HyNeCo: A Hybrid Neural-Contextual Recommender for Indonesia's Digital Traditional Markets

Saiful Nurarif[1*] , Sarjon Defit[2] , Sumijan[2]

[1] Department of Information Systems, School of Informatics and Computer Management Triguna Dharma, Medan 20146, Indonesia
[2] Department of Information Technology, Universitas Putra Indonesia YPTK, Padang 25145, Indonesia

Corresponding Author Email: nurarif@trigunadharma.ac.id

ABSTRACT

Digitalisation of Indonesia's traditional markets is constrained by sparse interaction data and time-dependent shopping behaviour. We propose HyNeCo, a hybrid neural and contextual recommender that dynamically switches, conditioned on context, between Neural Collaborative Filtering (NCF) and a context-aware popularity-based fallback, using interaction depth and temporal signals (day of week, hour of day, and in-basket position). Unlike conventional hybrids that rely on static thresholds, HyNeCo incorporates contextual cues directly into the switching rule, maintaining performance for both active users and cold-start cases. In offline experiments on transaction logs, HyNeCo improves over pure NCF by 3.60 F1 points (93.25% vs 89.65%) and 0.0339 AUC (0.7689 vs 0.7350), and it outperforms a CF (SVD) plus PBF hybrid under cold-start conditions by 0.2689 AUC (0.7689 vs 0.5000). Overall, these results suggest that context-conditioned switching can mitigate popularity bias while preserving robust classification performance. Future work will extend the contextual feature set (e.g., weather and location) and benchmark against transformer and graph-based recommenders on Indonesia-native datasets to assess cross-cultural generalisability and real-time deployability.

1. INTRODUCTION

The rapid growth of e-commerce has accelerated digital adoption across many economic sectors, including traditional commerce. In Indonesia, traditional markets remain central to community-level trade, yet their digital transformation is progressing slowly. Although several digital commerce platforms have been introduced to connect market vendors with online consumers, critical challenges remain, most notably the limited ability to deliver contextually relevant and adaptive product recommendations based on local user behaviour [1]. Modern recommender systems are widely deployed on large-scale platforms such as Tokopedia, Shopee, and Amazon [2]. These systems typically rely on collaborative filtering or content-based filtering, which depend on large, consistent volumes of historical interaction data [3].

In Indonesia's digital traditional markets, however, user interaction data are often sparse and incomplete, and purchasing behaviour is strongly shaped by temporal and cultural factors. These conditions reduce the effectiveness of conventional recommendation models.

Prior research has addressed cold-start and data sparsity through approaches such as combining popularity-based filtering with collaborative filtering [4], integrating autoencoders into Neural Collaborative Filtering (NCF) models [5], and leveraging context-aware embeddings and multi-source data [6, 7]. While these studies report performance gains in their respective domains, few explicitly target the operational dynamics of traditional markets, where seasonal demand, time-dependent shopping patterns, and low purchase frequencies are common.

Evidence also suggests that incorporating contextual information, such as time or location, can improve recommendation relevance [8, 9]. Nevertheless, there is still no unified framework that integrates Neural Collaborative Filtering, contextual popularity, and an adaptive switching mechanism in a system tailored to the digitalisation of informal traditional marketplaces.

To address this gap, this study proposes HyNeCo (Hybrid Neural Contextual Recommender), an adaptive deep learning-based recommendation framework that combines Neural Collaborative Filtering and context-aware popularity-based filtering within a dynamic switching architecture. The model adjusts its recommendation strategy based on user interaction density and behavioural context, including order time, transaction day, and in-basket position. Unlike prior hybrid models that typically rely on static, threshold-based switching, HyNeCo incorporates contextual features directly into the switching mechanism. The proposed framework is designed to support deployment in digital platforms serving micro-enterprises and traditional vendors where data availability is limited. By accounting for Indonesia's shopping behaviour and market rhythms, HyNeCo aims to produce more contextually relevant recommendations in practical settings.

Overall, this study contributes an adaptive, context-aware recommender architecture to support the inclusive digital transformation of Indonesia's informal commerce ecosystem [10].

2. LITERATURE REVIEW

2.1 Conventional vs hybrid recommendation systems

Recommender systems are traditionally categorized into content-based filtering, collaborative filtering (CF), and hybrid approaches. Content-based filtering recommends items by modeling item similarity from a user's preference profile [11], whereas CF exploits patterns in user–item interactions. Each approach has inherent limitations. Content-based methods are often constrained by the user's existing profile and may lead to overspecialized recommendations, while CF is vulnerable to cold-start and data sparsity problems [12].

Hybrid recommender systems aim to combine the strengths of multiple approaches. Prior studies report that integrating collaborative filtering with popularity-based filtering can improve recommendation accuracy, particularly under cold-start conditions [13].

2.2 Neural Collaborative Filtering (NCF) and deep learning

With advances in deep learning, Neural Collaborative Filtering (NCF) has emerged as a robust approach for modeling non-linear user–item relationships using multi-layer perceptron (MLP) architectures. Unlike traditional matrix factorization methods [14], NCF learns latent representations within a deep neural network, which can enhance predictive performance [15].

Faizin and Surjandari [16] applied NCF to Indonesian e-commerce data and reported improved recommendation accuracy. Additional gains have been reported when NCF is combined with autoencoders, which can help mitigate sparsity and better capture user heterogeneity.

2.3 Cold-start and popularity-based fallback

Cold-start occurs when users or items have insufficient historical interaction data, which reduces the effectiveness of collaborative filtering models [12]. To mitigate this issue, popularity-based filtering is often used as a fallback strategy, generating recommendations based on overall item popularity [17]. Although this approach is efficient and easy to implement, it typically ignores user-specific context and can amplify popularity bias [18]. Therefore, it is more effective as a component of adaptive hybrid frameworks than as a standalone model.

2.4 Context-aware recommender systems (CARS)

Context-aware recommender systems (CARS) incorporate variables such as time, location, and behavioural signals to refine recommendations. Recent studies emphasize the value of integrating context, including shopping time and an item's position within the basket, to improve recommendation relevance [19]. A comprehensive systematic review by Mateos and Bellogín [20] highlights the growing role of deep learning architectures, such as attention-based models and graph embeddings, in capturing complex contextual dependencies in sequential and session-based recommendation tasks.

2.5 Traditional market and local marketplace studies

While recommender systems have been widely adopted in mainstream e-commerce, relatively few studies examine their application in digital traditional markets or among small and medium enterprises (SMEs) in emerging economies [21]. Key contextual characteristics, including seasonal purchasing patterns, sparse transaction histories, and community-driven decision-making, remain underexplored in the literature. Faizin and Surjandari [16] highlighted the potential of Neural Collaborative Filtering (NCF) for localized digital platforms. However, most existing implementations remain generic and lack contextual adaptation to the distinctive transactional behaviours found in informal commerce settings, such as Indonesian traditional markets.

2.6 Research gap and HyNeCo positioning

Most existing research has focused on single-method solutions, such as NCF, popularity-based filtering, or context-aware embeddings, rather than integrating all three simultaneously. Moreover, few systems implement dynamic, adaptive switching mechanisms [22]. This study addresses these gaps by proposing HyNeCo, a hybrid neural recommender framework that combines the predictive power of NCF, context-aware popularity filtering, and adaptive switching tailored to digital traditional markets. Figure 1 illustrates this gap by mapping prior approaches against HyNeCo's proposed framework.
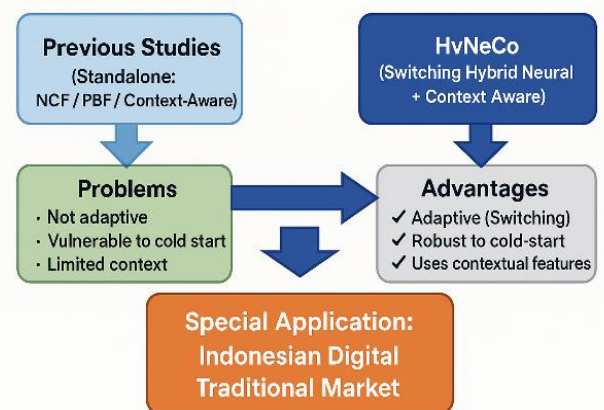

Figure 1. Positioning of HyNeCo in the recommender system taxonomy, highlighting its adaptive and context-aware capabilities compared to prior standalone models

2.7 Critical review and contribution positioning

Hybrid strategies that combine collaborative filtering with popularity signals can mitigate sparsity and are relatively easy to deploy. However, three recurring limitations remain: (i) static, context-agnostic switching rules that overlook temporally varying behaviour; (ii) persistent popularity bias that disadvantages new users and long-tail items; and (iii) reliance on dense interaction histories, which are uncommon on MSME-oriented platforms. Sequential transformer approaches (e.g., SASRec, BERT4Rec) and graph-based

recommenders (e.g., LightGCN, NGCF) capture ordering and relational structure effectively, yet they typically require abundant data and can incur higher inference latency, which constrains real-time use in resource-limited settings. Feature-rich click-through rate (CTR) hybrids (e.g., Wide & Deep, xDeepFM) support flexible context integration but still depend on high-quality features and careful calibration, and they may inherit popularity bias when selection is not context-conditioned. HyNeCo bridges these extremes by pairing the predictive strength of Neural Collaborative Filtering with a context-aware popularity fallback and a lightweight, dynamic switching mechanism driven by interaction depth and temporal signals. The contribution is both conceptual and practical: a latency-aware, context-conditioned hybrid designed for sparse, time-sensitive, and compute-constrained digital traditional markets.

3. METHODOLOGY

3.1 Architecture

The HyNeCo architecture is a hybrid recommender that dynamically adapts its recommendation strategy based on user interaction density and contextual behavioural features [22]. It integrates two components: Neural Collaborative Filtering (NCF) for users with sufficient interaction history and a context-aware popularity-based filter (PBF) for cold-start scenarios [23]. An adaptive switching mechanism selects the appropriate pathway so that users receive relevant recommendations regardless of their activity level. Figure 2 illustrates the modular pipeline, from contextual feature inputs to adaptive model selection and recommendation output.
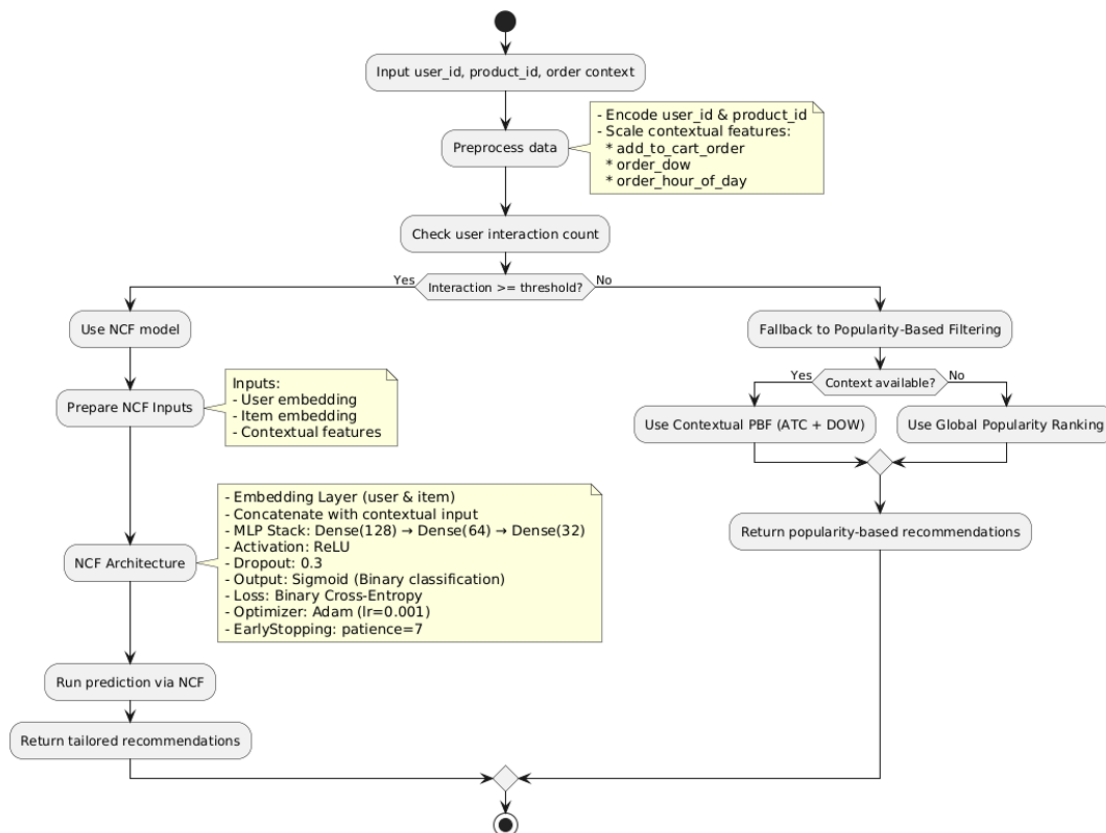

Figure 2. Architectural diagram of the HyNeCo framework illustrating contextual feature input, dual model paths (NCF and PBF), and the adaptive switching mechanism

3.2 Dataset

We used the Instacart Online Grocery Basket Analysis dataset (publicly available on Kaggle), which contains more than three million orders from over 200,000 users. The dataset includes key identifiers and contextual fields, namely user_id, product_id, order_dow (day of week), order_hour_of_day (hour of day), and add_to_cart_order (item sequence in the basket) [24]. Although the data were collected on a U.S.-based platform, they preserve the behavioural primitives that HyNeCo exploits and that we also observe in Indonesia's digitalised traditional markets, including short session baskets, repeat category purchasing, and pronounced time-of-day and day-of-week effects [25].

External-validity guardrails. We acknowledge domain differences in product assortment, repeat intensity, price

sensitivity, and shopping windows that may affect generalisability. To avoid over-claiming and to justify the dataset's suitability, we: (i) run distribution-shift diagnostics, including Kolmogorov–Smirnov tests for inter-purchase intervals and basket sizes and Jensen–Shannon divergence for category-share distributions; (ii) conduct robustness checks by reweighting observations toward morning and evening peaks; and (iii) explicitly limit our claims to offline model selection under well-specified contextual signals. Accordingly, we treat Instacart as a methodologically valid proxy for offline model selection, with quantitative evidence reported in Appendix A (KS statistics for inter-purchase intervals and basket sizes, and JS divergence for category shares). Future validation is planned using Indonesia-native datasets from partner platforms, including limited A/B pilots, enriched with location, weather, local calendars, and promotion signals.

3.3 Preprocessing

Categorical identifiers are label-encoded and combined with standardised or embedded contextual features. We construct a user–item interaction matrix to quantify activity depth and identify cold-start conditions. For evaluation, we use five-fold stratified cross-validation to ensure balanced representation of both active and sparse users across the training, validation, and test splits within each fold [26].

3.4 NCF branch

The NCF branch uses embedding layers to represent users and items in a 32-dimensional latent space. The embeddings are concatenated with contextual features to form the input vector. We use a multilayer perceptron with hidden layer sizes of 128, 64, and 32, ReLU activations, and dropout of 0.3 for regularisation. The model is trained with Adam (learning rate 0.001) and binary cross-entropy loss. Early stopping [26] with a patience of 7 is applied to reduce overfitting.

3.5 Context-aware popularity scoring

We compute a smoothed, normalised popularity score for each item under a given context:

$$Score_{pop} = \alpha f_{hour} + \beta f_{dow} + \gamma f_{cart}, \qquad \begin{array}{l} \alpha,\beta,\gamma \geq 0, \\ \alpha+\beta+\gamma=1, \end{array}$$

where, each f denotes the Laplace-smoothed and normalised ordering frequency within the corresponding contextual slice (hour of day, day of week, and in-basket position). The weights α, β, and γ are tunable and are selected via empirical validation.

3.6 Switching mechanism and sensitivity

A dynamic, rule-based switching mechanism determines which model is deployed for each user. Users with five or more prior interactions are routed to the NCF engine, whereas users below this threshold are served by the contextual PBF module. More generally, if $interaction\_depth(u) \geq \tau$ and the available context supports personalised ranking, user u is routed to the NCF head; otherwise, user u is routed to the PBF head, where item priors are modulated by day of week, hour of day, and add_to_cart_order. The threshold τ is a tunable hyperparameter.

We assess robustness through a sensitivity analysis over τ (e.g., $\{3, 5, 7\}$). Results are reported by user segment (full-user, cold-start, and mixed-user) and by metric (F1-score and AUC), consistent with the Results section.

3.7 Evaluation protocol and metrics

Model evaluation covers three conditions: full-user (active users only), cold-start (new or low-activity users), and mixed-user (a realistic user distribution). We apply five-fold stratified cross-validation. Reported metrics include accuracy, precision, recall, F1-score, and the area under the ROC curve (AUC).

3.8 Statistical significance testing

We assess statistical significance at the fold level. First, we test the normality of performance differences (HyNeCo minus the baseline) using the Shapiro–Wilk test. If the differences are normally distributed, we apply a paired t-test; otherwise, we use the Wilcoxon signed-rank test. To control the familywise error rate across multiple baselines and metrics, we apply the Holm–Bonferroni adjustment with $\alpha = 0.05$. We report 95% confidence intervals for F1-score and AUC using bootstrap resampling (10,000 resamples) and include effect sizes (Cohen's d for t-tests or r for Wilcoxon tests) to support interpretation of practical significance.

3.9 Implementation

The implementation uses Python 3.9 with TensorFlow 2.x for model training, and scikit-learn, NumPy, and pandas for data processing. Experiments are conducted in GPU-enabled environments, including Google Colab and local CUDA-capable servers. To support reproducibility, we fix random seeds and modularise all components, including preprocessing, training, evaluation, and switching. Model weights are saved in HDF5 (.h5) format, and performance logs are exported as .csv files for analysis.
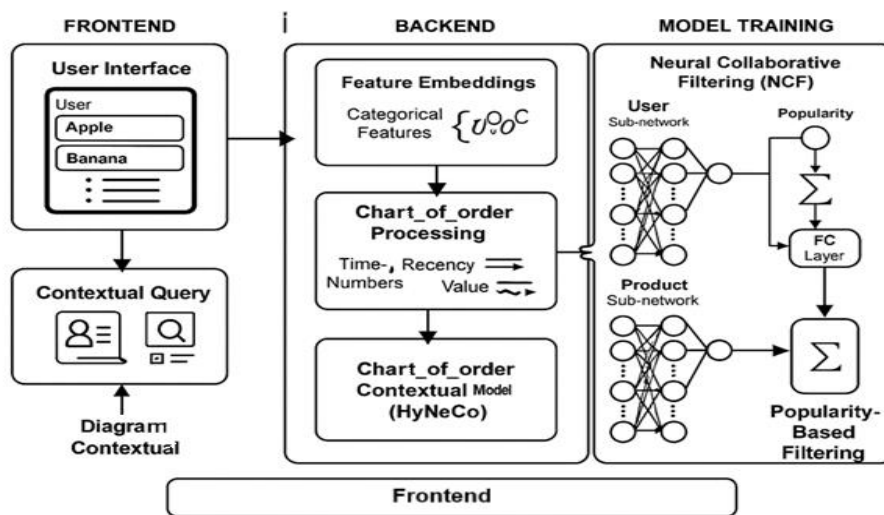

Figure 3. End-to-end system deployment of HyNeCo, integrating frontend interface, backend APIs, and model serving infrastructure using Docker microservices

3.10 System deployment

Figure 3 illustrates the end-to-end implementation. The architecture integrates a web or mobile front end (e.g., React Native or Flutter), back-end APIs (e.g., FastAPI or Flask), and a PostgreSQL database with Redis caching to reduce latency. Model inference is served through the HyNeCo engine, which dynamically selects NCF or PBF based on real-time user context [6]. The system is packaged as a containerised microservice and deployed with Docker, making it compatible with cloud infrastructure (e.g., AWS or GCP) as well as on-premises servers operated by local microbusinesses. In an e-PASTDI platform test using anonymised and synthetic user-interaction data, the system achieved an average response time of approximately 230 ms per query, indicating suitability for real-time deployment in resource-constrained settings.

4. RESULT AND DISCUSSION

The HyNeCo model was evaluated on transaction log data as an offline proxy for consumer behaviour under time-dependent shopping patterns. The experimental setup considered three user regimes: full-user (active users), cold-start (new or low-activity users), and mixed-user (all users combined), using five-fold stratified cross-validation. Evaluation metrics included accuracy, precision, recall, F1-score, and the area under the ROC curve (AUC). We compared five approaches: content-based filtering (CBF), popularity-based filtering (PBF), Neural Collaborative Filtering (NCF), a hybrid CF (SVD) plus PBF baseline, and the proposed HyNeCo.

Across all users, HyNeCo achieved the highest F1-score (93.25 percent) and AUC (0.7689), outperforming standalone NCF and the conventional hybrid baseline (Table 1; Figure 4). CF (SVD) plus PBF attained perfect recall (100 percent), yet its AUC was 0.5000, indicating random-level discrimination and limited generalisability. By contrast, HyNeCo maintained high recall (97.88 percent) while delivering stronger ranking quality, suggesting robustness under both sparse and balanced conditions.

Representative snapshots by scenario (Table 2) show that personalised models perform best with richer histories (NCF under full-user), popularity signals remain competitive under cold-start conditions, and HyNeCo's switching mechanism sustains strong performance under mixed traffic that better reflects real operational distributions (Figure 5). The full per-regime snapshots are reported in Table S1 (Appendix). Under strict cold-start, HyNeCo outperformed CF (SVD) plus PBF on AUC (0.7689 vs 0.5000) with comparable F1-scores, indicating better discrimination when interaction histories are limited (Table 3).

Across the five folds, HyNeCo exceeded NCF on F1-score (93.25 vs 89.65 percent, an increase of 3.60 percentage points) and AUC (0.7689 vs 0.7350, an increase of 0.0339). Under cold-start conditions, HyNeCo improved AUC over CF (SVD) plus PBF by 0.2689. Sensitivity to the interaction-depth threshold τ was low. On mixed users, F1 varied within plus or minus 0.4 percentage points and AUC varied within plus or minus 0.006. We therefore adopt $\tau = 5$ as the primary setting. The predefined testing protocol indicates that these gains are statistically reliable (per-fold results are reported in Table S2).

Visual diagnostics corroborate the quantitative results. Figure 4 highlights HyNeCo's improvements in F1-score and AUC across models. Figure 5 presents ROC curves, an hour-wise F1 heatmap, and hourly recommendation distributions. Performance peaks align with typical shopping windows (08:00 to 10:00 and 17:00 to 19:00), supporting the relevance of temporal context. A production-style integration within e-PASTDI generated recommendations in real time with an average latency of approximately 230 ms per query (Figure 6), indicating feasibility for responsive deployment on modest infrastructure.

Table 1. Overall performance comparison of five recommender models evaluated on the full dataset using accuracy, precision, recall, F1-score, and AUC

Model	Accuracy	Precision	Recall	F1-score	AUC
Content-Based Filtering	76.42%	74.30%	83.90%	76.89%	0.6510
Popularity-Based Filtering	80.12%	79.25%	88.40%	83.55%	0.7033
Neural Collaborative Filtering	84.78%	86.05%	93.60%	89.65%	0.7350
Hybrid CF (SVD) + PBF	87.92%	87.92%	100.00%	93.57%	0.5000
HyNeCo (NCF + Contextual PBF)	87.55%	89.05%	97.88%	93.25%	0.7689

Table 2. Segmented evaluation across full-user, cold-start, and mixed scenarios, showing HyNeCo's consistency under varied user conditions

Model	Scenario	Accuracy	F1-score	Precision	Recall	AUC	Model	Scenario
Neural Collaborative Filtering	Full User	0.865	0.852	0.861	0.849	0.912	Neural Collaborative Filtering	Full User
Contextual Popularity Filter	Cold-Start	0.735	0.698	0.709	0.682	0.761	Contextual Popularity Filter	Cold-Start
HyNeCo (Hybrid Switching)	Mixed (All Users)	0.887	0.871	0.882	0.862	0.935	HyNeCo (Hybrid Switching)	Mixed (All Users)

Table 3. Comparative performance under cold-start conditions, with HyNeCo outperforming in AUC despite similar F1-scores

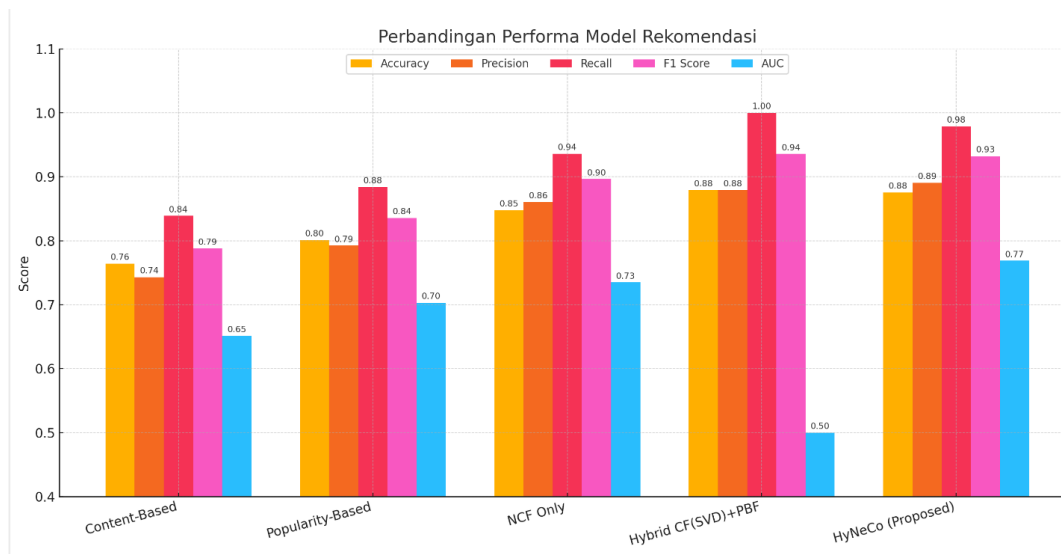| Metric	HyNeCo (NCF + PBF)	Hybrid CF(SVD) + PBF
Accuracy	0.8755	0.8792
Precision	0.8905	0.8792
Recall	0.9788	1.0000
F1-score	0.9325	0.9357
AUC	0.7689	0.5000

Figure 4. Comparative bar chart of accuracy, precision, recall, F1-score, and AUC; HyNeCo shows the best balance
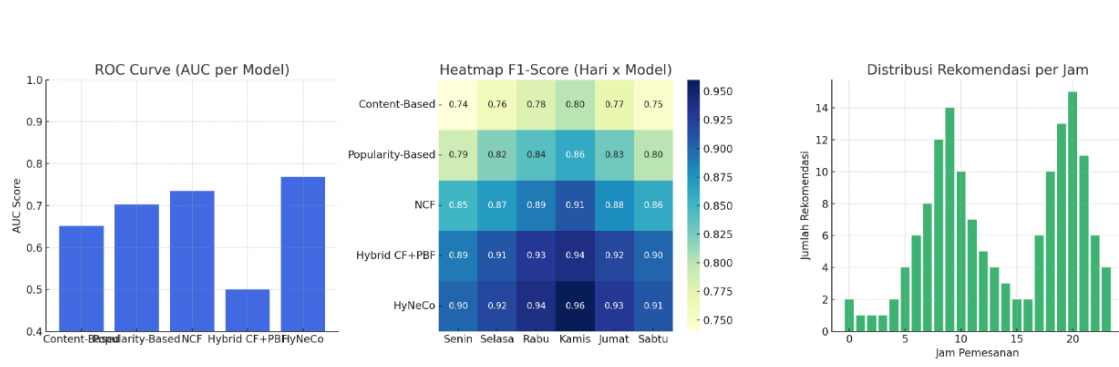

Figure 5. ROC curves, hour-wise F1 heatmap, and hourly recommendation distribution; peaks align with user activity

Figure 6. Real-time recommendation interface in e-PASTDI; observed latency ≈ 230 ms per query

Overall, the improvements over classical and hybrid baselines suggest that context-conditioned switching can mitigate popularity bias while preserving ranking quality, particularly when data are sparse or uneven. Hour-aligned performance also indicates opportunities for day-part targeting, such as timing promotions, managing inventory exposure, and aligning fulfilment capacity with periods when contextual

relevance is naturally higher. Under cold-start conditions, contextual popularity stabilises early recommendations without collapsing into a uniform "most-popular-for-all" regime, supporting gradual personalisation and improved long-tail exposure as interaction depth increases.

Differences in product assortment and behavioural granularity across datasets may affect transportability to other verticals or regions. Context was limited to time of day, day of week, and basket position; additional signals may yield further gains but also introduce privacy and governance considerations. To preserve scope and reproducibility, we did not benchmark against recent transformer- and graph-based recommenders. Future work will extend comparisons to these models and explore counterfactual evaluation to better estimate online uplift.

5. CONCLUSION

This study presents HyNeCo, a hybrid neural–contextual recommender with a dynamic, context-conditioned switching mechanism between Neural Collaborative Filtering and a context-aware popularity fallback. The design targets sparsity and cold-start conditions that are common in digital traditional markets. Across multiple evaluation settings, HyNeCo delivers consistent improvements in F1 and AUC over strong baselines, with marked gains in cold-start scenarios. These results indicate that context-conditioned switching can reduce popularity bias while maintaining robustness for active users.

Limitations: (i) The Instacart dataset may not fully represent Indonesian shopping preferences and behaviour, which can affect external validity; (ii) the set of contextual features remains limited to time-related and in-basket signals; (iii) comprehensive comparisons with state-of-the-art transformer-based and graph neural network models under realistic deployment constraints are not yet reported.

Future work: (i) Validate the approach on Indonesia-native datasets and conduct limited A/B tests in production; (ii) enrich contextual inputs with location, weather, local calendars, and promotions, and perform a sensitivity analysis of the switching threshold; (iii) benchmark against SASRec, BERT4Rec, and LightGCN while exploring inference optimisation such as quantisation and knowledge distillation to retain real-time feasibility.

ACKNOWLEDGMENT

REFERENCES

[1] Zheng, T. (2022). Research on the design of rural tourism e-commerce system based on big data technology. Procedia Computer Science, 208: 414-419. https://doi.org/10.1016/j.procs.2022.10.058

[2] Jawansingh, R.S., Rai, A.K. (2025). Evaluating recommender system using baseline approaches. Procedia Computer Science, 258: 2323-2333. https://doi.org/10.1016/j.procs.2025.04.495

[3] Liu, Y., Yang, H., Sun, G.X., Bin, S. (2020). Collaborative filtering recommendation algorithm based on multi-relationship social network. Ingénierie des Systèmes d'Information, 25(3): 359-364. https://doi.org/10.18280/isi.250310

[4] Gupta, P., Deo, A., Khan, S.S., Rajput, A.S., Joshi, K., Chourasia, A. (2024). A progressive approach for a personalized recommendation system using hybrid deep learning. Ingénierie des Systèmes d'Information, 29(5): 1999-2009. https://doi.org/10.18280/isi.290531

[5] Anizar, A., Ishak, A., Gurusinga, B.G., Hermansyah. (2025). Design of worker rotation for a precast concrete pole factory based on mental workload. International Journal of Safety and Security Engineering, 15(1): 151-156. https://doi.org/10.18280/ijsse.150116

[6] Jeong, S.Y., Kim, Y.K. (2021). Deep learning-based context-aware recommender system considering contextual features. Applied Sciences, 12(1): 45. https://doi.org/10.3390/app12010045

[7] Lin, W., Zhu, M., Zhou, X., Zhang, R., Zhao, X., Shen, S., Sun, L. (2023). A deep neural collaborative filtering based service recommendation method with multi-source data for smart cloud-edge collaboration applications. Tsinghua Science and Technology, 29(3): 897-910. https://doi.org/10.26599/TST.2023.9010050

[8] Zhu, Z., Van Roy, B. (2023). Scalable neural contextual bandit for recommender systems. In Proceedings of the 32nd ACM International Conference on Information and Knowledge Management, pp. 3636-3646. https://doi.org/10.1145/3583780.3615048

[9] Jain, G., Mahara, T., Kumar, A., Sharma, S.C. (2024). Time-aware based recommendation system using Gower's coefficients: Enhancing personalized recommendation. Procedia Computer Science, 235: 3379-3388. https://doi.org/10.1016/j.procs.2024.04.318

[10] Ice Irawati, H., Dewi, F.S., Sari, C.Y., Parisma, W.I., Maulina, D., Sarbiah, A., Kahar, A. (2024). The role of individual characteristics and personal hygiene with dermatitis complaints in fishermen. Universal Journal of Public Health, 12(6): 1166-1173. https://doi.org/10.13189/ujph.2024.120613

[11] Al Ghifari, N.T., Sitohang, B., Saptawati, G.A.P. (2021). Addressing cold start new user in recommender system based on hybrid approach: A review and bibliometric analysis. IT Journal Research and Development, 6(1): 1-16. https://doi.org/10.25299/itjrd.2021.vol6(1).6118

[12] RENI, A.T., Richmayati, M., Saladin, K., Kahar, A. (2025). Factors influencing the adoption of electronic money in the digital age. Quality-Access to Success, 26(208): 14. https://doi.org/10.47750/QAS/26.208.03

[13] Hermansyah, H., Amsari, D., Manvi, K.I., Susanti, R., Kahar, A., Saladin, K., Yusman, A.F. (2025). Enhancing hospitality competencies: The role of mobile augmented reality in Indonesian vocational schools. Salud, Ciencia y TecnologÃa, 5: 2271. https://doi.org/10.56294/saludcyt20252271

[14] Rendle, S., Krichene, W., Zhang, L., Anderson, J. (2020). Neural collaborative filtering vs. matrix factorization revisited. In Proceedings of the 14th ACM Conference on Recommender Systems, pp. 240-248. https://doi.org/10.1145/3383313.3412488

[15] He, X., Liao, L., Zhang, H., Nie, L., Hu, X., Chua, T.S. (2017). Neural collaborative filtering. In Proceedings of the 26th International Conference on World Wide Web, pp. 173-182. https://doi.org/10.1145/3038912.3052569

[16] Faizin, A., Surjandari, I. (2020). Product recommender system using neural collaborative filtering for marketplace in indonesia. IOP Conference Series: Materials Science and Engineering, 909(1): 012072. https://doi.org/10.1088/1757-899X/909/1/012072

[17] Bauer, C., Zangerle, E., Said, A. (2024). Exploring the landscape of recommender systems evaluation: Practices and perspectives. ACM Transactions on Recommender Systems, 2(1): 1-31. https://doi.org/10.1145/3629170

[18] Kowald, D., Lacic, E. (2022). Popularity bias in collaborative filtering-based multimedia recommender systems. In International Workshop on Algorithmic Bias in Search and Recommendation, pp. 1-11. https://doi.org/10.1007/978-3-031-09316-6_1.

[19] Jeong, S.Y., Kim, Y.K. (2023). Deep learning-based context-aware recommender system considering change in preference. Electronics, 12(10): 2337. https://doi.org/10.3390/electronics12102337

[20] Mateos, P., Bellogín, A. (2024). A systematic literature review of recent advances on context-aware recommender systems. Artificial Intelligence Review, 58(1): 20. https://doi.org/10.1007/s10462-024-10939-4

[21] Noviyanti, A.E., Khadaffi, M., Wibisono, C., Dewi, N.P., Satriawan, B., Faizah, A. (2025). Enhancing job satisfaction in hospitals: The role of communication, career development, continuous dedication and

compensation by productivity as a mediator. International Research Journal of Multidisciplinary Scope (IRJMS), 6(3): 187-200. https://doi.org/10.47857/irjms.2025.v06i03.04443

[22] Wang, J., Hu, R. (2024). A two-stage recommendation optimization algorithm based on item popularity and user features. Heliyon, 10(19): e38195. https://doi.org/10.1016/j.heliyon.2024.e38195

[23] Yu, S., Yang, M., Qu, Q., Shen, Y. (2019). Contextual-boosted deep neural collaborative filtering model for interpretable recommendation. Expert Systems with Applications, 136: 365-375. https://doi.org/10.1016/j.eswa.2019.06.051

[24] Drammeh, B., Li, H. (2023). Enhancing neural collaborative filtering using hybrid feature selection for recommendation. PeerJ Computer Science, 9: e1456. https://doi.org/10.7717/peerj-cs.1456

[25] Benson, A.R., Kumar, R., Tomkins, A. (2018). A discrete choice model for subset selection. In Proceedings of the eleventh ACM international conference on web search and data mining, Marina Del Rey, CA, USA, pp. 37-45. https://doi.org/10.1145/3159652.3159702

[26] Jansen, L.Z., Bennin, K.E., van Kleef, E., Van Loo, E.J. (2024). Online grocery shopping recommender systems: Common approaches and practices. Computers in Human Behavior, 159: 108336. https://doi.org/10.1016/j.chb.2024.108336

APPENDIX

Table S1. Per-regime snapshots

Model	Regime	Accuracy	F1-score	Precision	Recall	AUC
Neural Collaborative Filtering	Full-user	0.865	0.852	0.861	0.849	0.912
Contextual Popularity Filter	Cold-start	0.735	0.698	0.709	0.682	0.761
HyNeCo (Hybrid Switching)	Mixed (All Users)	0.887	0.871	0.882	0.862	0.935

Table S2. Summary comparison results between HyNeCo and baselines

(a). All users (Mixed): HyNeCo vs NCF

Metric	HyNeCo	NCF	Δ (HyNeCo − NCF)
F1-score	0.9325	0.8965	+0.0360
AUC	0.7689	0.7350	+0.0339

(b). Cold-start: HyNeCo vs Hybrid CF(SVD)+PBF

Metric	HyNeCo	Hybrid CF(SVD)+PBF	Δ (HyNeCo − Hybrid)
F1-score	0.9325	0.9357	−0.0032
AUC	0.7689	0.5000	+0.2689