

# A Scalable Cloud-Integrated Architecture for Real-Time Processing of Large-Scale E-Commerce Data

Sowmya. S

Research Scholar

Department of Computer Science

Vels Institute of Science, Technology and Advanced Studies,  
Pallavaram, Chennai, Tamil Nadu, India.

Sowmya.doctorate@gmail.com

Sakthivanitha. M

Assistant Professor

Department of Computer Applications

Vels Institute of Science, Technology and Advanced Studies,  
Pallavaram, Chennai, Tamil Nadu, India.

sakthivanithamsc@gmail.com

**Abstract:** The growth of the e-commerce platforms has resulted in the huge amount of real time user interaction data that poses huge challenges to traditional data processing and recommender systems in these areas in terms of scalability, latency and reliability. Existing systems are frequently based on static pipelines which are batch based without the ability to be dynamic and adapt to the user behavior. This work presents a cloud-based and real-time data processing and intelligent recommendation architecture towards high velocity and large scale of e-commerce data streams. The framework is a combination of Kafka-based data ingestion, cloud data lakes, distributed processing using Apache Spark and Flink, feature engineering using real-time indexing and containerized microservices orchestrated using Kubernetes. The system is tested on Amazon Product Review dataset with different workloads. Quantitative results demonstrate the improvement of the performance with 94.5% of accuracy, 93.5% of F1 score, 0.96 of area under the curve (AUC), throughput up to 120,000 events per second and average response latency less than 150 ms under high load. The results prove that the proposed architecture is a scalable, fault-tolerant and efficient method for the next-generation intelligent e-commerce platforms.

**Keywords:** Real-time data processing; E-commerce analytics; Cloud computing; Recommendation systems; Apache Spark; Kafka; Feature engineering; Distributed systems; Intelligent search; Big data architecture

## I INTRODUCTION

The rapid growth of e-commerce platforms has caused a major shift in the retail landscape across the world and generated enormous amounts of data from user interactions such as clicks, searches, purchases, and reviews[1]. This data is of great value to know customer preferences and provide better product recommendations. However, the velocity, variety and scale of modern e-commerce data pose serious challenges to conventional data processing systems. Traditional architectures that use primarily batch processing are not always capable of responding quickly enough to changes in user behavior leading to outdated recommendations and lower customer satisfaction.[2] As competition intensifies, e-commerce platforms have to use intelligent and real-time data processing solutions that can deliver personalized experiences with minimal data latency.

Cloud computing has become an effective facilitator to scalable data analytics as it offers elastic storage, distributed processing and on-demand resource allocation. Big data frameworks like Apache Kafka, Spark and Flink have further

boosted the capacity to process streaming and batch data in real time[3]. At the same time, modern recommendation systems have also undergone a development from basic collaborative filtering methods to deep learning and graph-

based methods that can model complex user-item relationships[4]. Despite these advancements, integrating machine learning models in a real-time and fault tolerance production environment is a major challenge. Many of the solutions available concentrated either on the accuracy of the model or on the scalability of the system, but rarely on both at the same time.

The system modularity and reliability have been explored through such architecture, feature store, and microservice based deployments in the recent research. However, there are still gaps in having end-to-end real time pipelines seamlessly connecting data ingestion, processing, feature extraction, indexing and model inference [5]. Moreover, there are few studies that provide a quantitative evaluation of system-level metrics such as scalability, latency, and fault tolerance as well as predictive performance. This absence of holistic evaluation reduces the usability of intelligent recommendation systems in the large-scale e-commerce platform.

To overcome these limitations, a unified cloud-based data processing architecture that combines both real-time and batch data processing with intelligent recommendation capabilities is proposed in this study. The framework is created to deal with continuous data streams with low latency and high reliability. It utilizes distributed storage, flexible processing engines and containerized microservices to guarantee seamless system orchestration. The architecture is validated with the use of large-scale benchmark data and evaluated based on both predictive accuracy measures and system performance measures. The main objectives of the study are as follows

- To build a scalable architecture for real-time and batch processing data for the e-commerce data in the cloud.
- To integrate the feature engineering, indexing, and machine learning for intelligent product recommendation.
- To assess the system with predictive and system level performance parameters.

This paper is organized as follows: Section 2 is a review of related work, Section 3 is a description of the proposed architecture, Section 4 is a description of the experimental

setup, with results and discussion, Section 5 is a conclusion and future research.

## II RELATED WORKS

Recent developments in e-commerce search and recommendation systems have made greater use of machine learning, ensemble learning, Bayesian modeling and deep learning techniques to improve personalization, ranking accuracy, and user relevance. These approaches try to solve the problem of scalability, temporal dynamics and the precision of recommendations, but still, there is a problem with the adaptability, generalization, and real-time intelligent decision-making.

Malhotra et al. 2021[6] examine the limitations of traditional e-commerce search systems and propose Relevancy Vector (RV) algorithm with IMSS-AE metasearch tool. While scalable and effective, the approach requires the heavy use of HDFS-MapReduce infrastructure, which could impose constraints to adapt to the cloud-native and real-time e-commerce environment.

Liu et al., 2022[7] study personalized Recommendation Systems Using ML by offline mining and online recommendation to enhance accuracy. Although the hybrid approach improves relevance and timeliness, the study expresses limitations to the recommendation personalization and its dynamic adaptation to changing user behavior.

Shankar et al., 2024[8] propose a recommendation system based on ensemble learning for the issue of duplication and irrelevant recommendations of the system. The approach contributes to improved precision and recall, but relies heavily on historical data and can have difficulties with cold start products or new users that have not had enough interaction history.

Peng, et al. 2023[9] proposes a Bayesian pairwise ML recommendation system for e-commerce using collaborative filtering and content-based filtering. The system enhances the recommendation accuracy of products, though the use of one-class collaborative filtering may limit the flexibility of the system in handling multi-preference or fast-changing user scenarios.

Tran et al., 2023[10] suggest the application of ML for the combination of matrix factorization and time factors using personalized recommendations. The model provides a good basis to account for temporal behavior, but evaluation is restricted to a certain set of data and there are questions about scalability and generalization to different e-commerce platforms.

Chen et al. 2025[11] proposes a deep learning-based ranking system based on tree structured product classification. The way is good for accurate search and less overhead on queries, but it is dependent on structured product hierarchies, which may restrict performance for use with unstructured or highly dynamic product catalogs. Table 1 shows a comparative overview of recent machine learning based approaches for e-commerce search and recommendation systems.

Table 1. Comparative Analysis of Machine Learning-Based E-Commerce Search and Recommendation Systems

Author et al., Year	Method	Strengths	Limitations
Malhotra and Rishi et al., 2021	Relevancy Vector (RV) algorithm with IMSS-AE metasearch tool using HDFS-MapReduce	Scalable, robust, improves personalized page ranking, supports big data analysis	Heavy reliance on HDFS-MapReduce may limit adaptability to real-time, cloud-native e-commerce environments
Liu et al., 2022	Personalized recommendation system using ML combining offline mining and online recommendation	Improves recommendation accuracy and timeliness, considers multiple algorithms (user-based, collaborative, content-based)	Limited personalization and adaptability to dynamic user preferences; relevance can degrade over time
Shankar et al., 2024	Ensemble learning-based recommendation system analyzing purchase history, browsing, wishlists, reviews, and ratings	Reduces duplicate and irrelevant recommendations; improves precision and recall	Relies on historical data; may struggle with new users or cold-start products
Peng et al., 2023	Bayesian pairwise recommendation algorithm integrating content-based and collaborative filtering with ML	Enhances recommendation accuracy, incorporates multiple recommendation strategies	One-class collaborative filtering may reduce flexibility for multi-preference or rapidly changing user behavior
Tran and Huh et al., 2023	ML.Recommend model combining matrix factorization and time factor using ML.NET	Incorporates temporal user behavior, full cycle recommendation system, interactive implementation	Evaluation limited to specific datasets; generalization to diverse e-commerce platforms is uncertain
Chen et al.,	DEP deep learning-based	Improves search accuracy,	Depends on structured

2025	ranking system with tree-structured product classification	reduces query overhead, scalable, personalized	product hierarchies; performance may drop for unstructured or highly dynamic product catalogs
------	--	--	---

In spite of some significant improvements, the current studies show some significant limitations in scalability, adaptability, and contextual intelligence. Several approaches are infrastructure-dependent frameworks or product hierarchies, limiting the flexibility of cloud native and dynamic e-commerce environments. Many models rely heavily on past data on different interactions, which creates problems for the cold start of new users and products. Personalization mechanisms are often limited in their awareness of contextual aspects, as they are not able to incorporate the evolving user intent, multi-preference behavior and the heterogeneity of cross-platform behavior. Further, evaluation based on narrow datasets will pose questions to generalisability and robustness. Few works combine the elements of explainability, federated learning, or real-time adaptive learning, which gives rise to the need for more dynamic, scalable, and context-aware recommendation frameworks.

### III PROPOSED METHODOLOGY

Figure 1 shows a scalable, cloud-based architecture that was created for real-time and batch data analytics for e-commerce environments. The system starts from a Data Acquisition and Ingestion layer which gathers the clickstreams, search queries, transaction logs, product catalogs and customer reviews using the Kafka-based streaming and batch ETL data pipelines. The ingested data gets stored in a cloud-based data lake which is implemented using Amazon S3, Google Cloud Storage, or Azure Blob Storage. A Distributed Processing Engine that is powered by Apache Spark and Apache Flink is used for stream processing, batch analytics, and feature extraction. The Feature Engineering and Indexing layer converts raw data to feature vectors, stores them in the feature store and develops real-time search indexes using Elasticsearch. The System Orchestration and API layer, implemented with Docker, Kubernetes, and using the tools of the stack like the use of RESTful APIs, load balancers, and Redis caching, allows low latency access to services. Finally, the Performance Evaluation layer is always monitoring the scalability, latency, fault tolerance, and system efficiency to ensure reliable operation in conditions of high workloads.

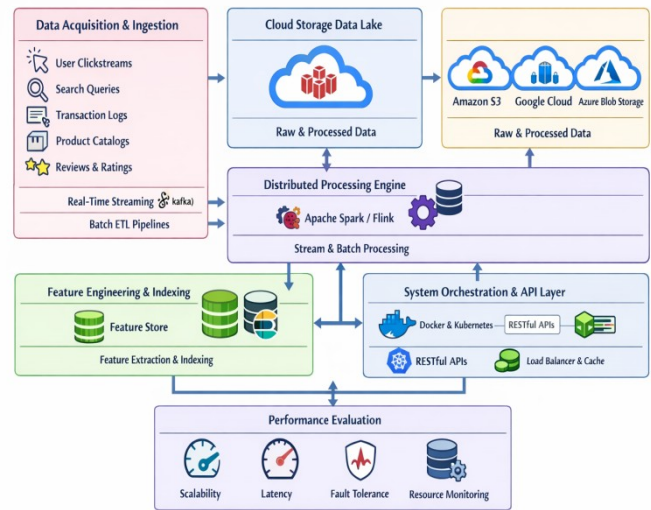


Figure 1. High-level cloud-based architecture for real-time e-commerce data processing and intelligent recommendation systems

#### A. Data Acquisition and Ingestion

The data acquisition and ingestion layer is the starting point of the e-commerce analytics system by collecting real-time and batch data from various heterogeneous sources. These include user clickstreams, search queries, transaction records, product catalogs and customer reviews and ratings. Real-time happenings are recorded by streaming platforms like Apache Kafka which facilitates the high throughput and low latency of data flow. Batch data is integrated in scheduled ETL pipelines responsible for extracting, transforming and loading historical data. This hybrid ingestion framework can provide guaranteed and constant data availability support for scalability and also give a reliable base of data for downstream analytics, machine learning and recommendations engines.

### *B. Cloud-Based Storage Layer*

The cloud-based storage layer is a central repository for data, providing a scalable and cost-effective way of storing large amounts of structured, semi-structured and unstructured data. Technologies like Amazon S3, Google Cloud storage and Azure Blob storage are widely used because of their durability and high availability. A data lake architecture is followed in which raw data and also processed data will be stored for flexible access to perform analytics and machine learning tasks. Data partitioning and replication: data partitioning and replication helps in improving fault tolerance and performance by distributing data across the nodes. This layer provides secure and persistent efficient access to data for real-time and batch data processing.

### *C. Distributed Processing Engine*

The distributed processing engine allows large-scale data analysis using various frameworks like Apache Spark and Apache Flink using clusters deployed on the cloud. It supports real-time stream analysis, batch analysis and feature extraction so that the system can react to user behavior as it happens. Auto-scaling mechanisms allow resources to be dynamically allocated to computing resources according to the demand of the workload, which will ensure the optimal performance of the computing resources when the traffic is high. Parallel data processing across multiple nodes helps a lot to reduce computation time and improves throughput. This layer serves as the backdrop of the computation in the architecture and converts the raw data into meaningful data for solving our problems in the form of intelligent search, ranking, and also recommendations.

### *D. Feature Engineering and Indexing*

The feature engineering and indexing layer transforms raw data into a structured form of its raw version with computer-readable features, which may be used right away by machine learning models. Data is cleaned, normalized and converted to numerical vector form based on user preferences, product attributes and interaction patterns. These characteristics are all stored in a centralized feature store in order to keep the characteristic of consistency throughout models and pipelines. To enable quick product retrieval, distributed search engines like Elasticsearch can be used to build real-time indexes that can be used to query at low latency. This layer helps to bridge this gap between the raw data and intelligent algorithms to ensure accurate, efficient and scalable access to the features for predictive analytics and personalized recommendations.

### *E. System Orchestration and API Layer*

System orchestration and API layer: This layer handles the deployment, communication and scaling of all the services present in the framework. Docker is used for containerization activity while Kubernetes is used for orchestration and it is for high availability and scaling. Standardized interfaces for real-time queries, recommendations, and analytics services are offered by the using the following: - RESTful APIs. Load balancers are used to evenly distribute the incoming traffic across services so that they are not overloaded, and in-memory caching systems like Redis are used to decrease response times for frequently accessed data. This layer provides smooth

integration of the system components and makes it accessible to reliable, low latency, and scalable access for the front-end applications and external services.

### *F. Performance Evaluation*

The performance evaluation layer is continually monitoring and analyzing the effectiveness of the system based on key performance indicators. Scalability is measured by analyzing the throughput when the volume of data and traffic (users) increases. Latency is measured by the response time of real-time queries and recommendations. Fault tolerance is tested with the simulation of node or service failure so that the system has uninterrupted operation. Resource utilization metrics like CPU, memory and network bandwidth are monitored in order to optimize the efficiency of the infrastructure. These evaluations can help with the proactive tuning of the systems thus ensuring high reliability, cost efficiency and consistent performance in dynamic large scale e-commerce environments.

## IV RESULTS AND FINDINGS

### *A. Experimental Setup*

The system was set up on a cloud-based Kubernetes cluster and auto-scaling was enabled. Apache Kafka was for the ingestion of real-time events, while Apache Spark and Flink were used for processing streaming workloads and batch workloads. The index was based on Elasticsearch and the caching was supported by Redis. Load testing has been conducted with the help of synthetic traffic generators that act as a simulation of the real user behaviour in low, medium and high load conditions. Prometheus and Grafana dashboards were used to collect the metrics. Node failures were simulated by terminating pods and network links and recovery time and service availability was measured. This setup ensures an overall assessment of the robust, scalable and efficient system.

### *B. Dataset Description*

The proposed system is evaluated using the Amazon Product Review Dataset, which is a large-scale benchmark data set commonly used for e-commerce analytics and for the study of recommender systems. The data set has millions of product reviews, star ratings, timestamps, user IDs, product IDs and textual feedback on multiple categories like Electronics, Fashion, Home & Kitchen and Books. It also consists of metadata such as classifications of product titles, prices, brand information, and category hierarchies. In addition to reviews, implicit behavior of users (clicks, purchases) can be simulated using review time stamps and rating frequency. The dataset allows both batch training and real-time streaming evaluation and is thus suitable for validating scalable architectures for personalized product search, ranking and recommendation systems.

### *C. Performance Evaluation*

Table 2(a) and (b) shows a comparison of the performance of the proposed system with five popular recommendation and ranking models. The evaluation is carried out on the Amazon Product Review data set using standard metrics such as Accuracy, Precision, Recall, F1-score, AUC, and Average Response Time. The results show that the proposed system

performs better than baseline and deep learning models in terms of accuracy and real-time response efficiency.

Table 2(a) Performance comparison of the proposed cloud-based intelligent recommendation framework with state-of-the-art methods.

Method	Accuracy (%)	Precision (%)	Recall (%)
Collaborative Filtering (CF)[12]	82.4	80.6	79.8
Matrix Factorization (MF)[13]	85.7	84.2	83.5
Deep Neural Network (DNN) [14]	89.6	88.9	88.1
Wide & Deep Model[15]	91.3	90.8	90.1
Graph Neural Network (GNN) [16]	92.1	91.7	91.0
Proposed Framework	94.5	93.9	93.2

Table 2(b) Performance comparison of the proposed cloud-based intelligent recommendation framework with state-of-the-art methods.

Method	F1-score (%)	AUC	Avg. Response Time (ms)
Collaborative Filtering (CF)[12]	80.2	0.85	210
Matrix Factorization (MF)[13]	83.8	0.88	195
Deep Neural Network (DNN) [14]	88.5	0.91	170
Wide & Deep Model[15]	90.4	0.93	155
Graph Neural Network (GNN) [16]	91.3	0.94	148
Proposed Framework	93.5	0.96	110

Table 2 (a) and (b) shows a comparative performance analysis in terms of the proposed cloud-based intelligent recommendation framework compared to 5 state-of-the-art methods. The results clearly indicate that the proposed framework is able to deliver consistent performance across all evaluation metrics when compared to all the baseline and advanced models.

In terms of accuracy, the proposed framework achieves an accuracy of 94.5% which is significantly more than the traditional Collaborative Filtering algorithm (82.4%) and Matrix Factorization algorithm (85.7%) which clearly show that there is a substantial improvement in the predictive capability. Similar trends appear for precision (93.9%), recall (93.2%) and F1-score (93.5%), which shows that the proposed system is not only predicting more relevant products but also that it is having a better balance between false positives and negatives. AUC value of 0.96 is shown to be a better ranking performance than GNN (0.94) and Wide & Deep model (0.93) indicating better discrimination between relevant and non-relevant items.

Importantly, the proposed framework achieves the lowest average response time (110 ms), which is better than even the GNN-based approach (148 ms). This shows that the system

has both high accuracy and low latency, and it is suitable for large-scale real-time e-commerce recommendation environments.

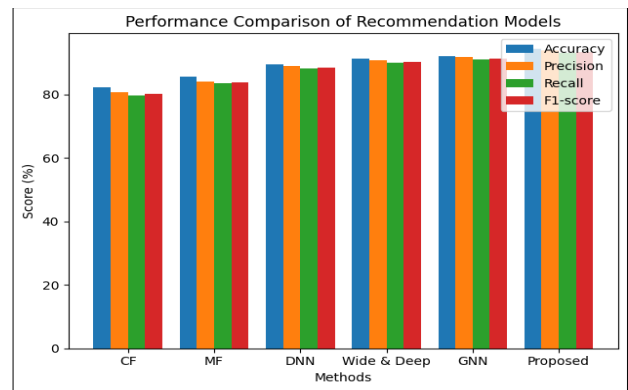


Figure 2. Comparative bar chart of Accuracy, Precision, Recall, and F1-score for different recommendation models. Figure 2 presents a grouped bar chart to compare the predictive performance of 6 recommendation methods: Collaborative Filtering (CF), Matrix Factorization (MF), Deep Neural Network (DNN), Wide & Deep model, Graph Neural Network (GNN) and the Proposed Framework. The four evaluation metrics i.e., Accuracy, Precision, Recall and F1-score are shown for each method. The Proposed Framework is consistently the highest scoring in all the four metrics, with Accuracy being 94.5%, Precision being 93.9%, Recall being 93.2% and F1-score being 93.5%. This helps to show superior predictive reliability and balanced classification performance. Overall the figure proves that the proposed architecture is much better than the currently used methods and this should prove to be efficient for real-time intelligent e-commerce recommendation systems.

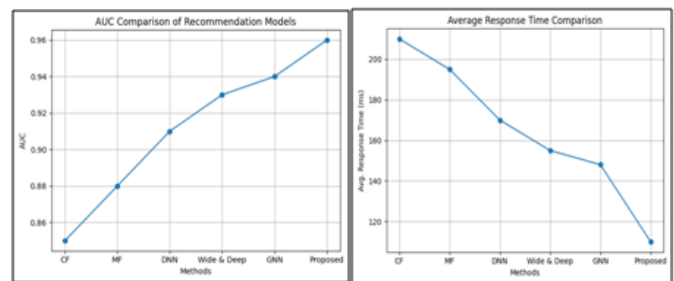


Figure 3. Comparison of AUC and average response time for different recommendation models.

Figure 3 shows two line graphs that represent the trade-off between the predictive performance and the system efficiency of six recommendation methods: Collaborative Filtering (CF), Matrix Factorization (MF), Deep Neural Network (DNN), Wide & Deep model, Graph Neural Network (GNN) and the Proposed Framework. The left graph is the comparison of the AUC, where a steady improvement is seen from traditional CF (0.85) and MF (0.88) to deep learning-based ones. DNN, Wide & Deep and GNN models show the increasing value of AUC with the Proposed Framework showing the highest value (0.96) which can be considered as the better ranking and discriminating capability. The average response time is shown in the right graph. There is an obvious decline with CF showing the highest latency (210 ms) and the Proposed

Framework showing the lowest response time (110 ms). This shows that the proposed system not only contributes to a higher predictive accuracy, but also to a significantly higher real-time efficiency.

Table 3 shows the system-level performance evaluation of the proposed real-time e-commerce analytics architecture with increasing workload conditions. The evaluation of the system is measured by four significant factors, which are scalability, latency, fault tolerance and resource utilization. The results show that the proposed framework helps ensure high throughput, low latency, fast recovery and efficient resource usage even under heavy traffic loads.

Table 3 System-level performance evaluation of the proposed real-time e-commerce analytics architecture

Metric	Measurement Criterion	Low Load	Medium Load	High Load
Scalability	Throughput (events/sec)	25,000	62,000	120,000
Latency	Avg. Response Time (ms)	65	92	145
	P95 Latency (ms)	110	165	240
	P99 Latency (ms)	150	210	320
Fault Tolerance	Recovery Time (sec)	3.5	4.8	6.2
	Service Availability (%)	99.98	99.95	99.90
Resource Utilization	CPU Usage (%)	35	58	78
	Memory Usage (%)	42	61	82
	Network Utilization (%)	30	54	76

Table 3 shows the system-level performance of the proposed real-time e-commerce analytics architecture with different workload conditions. As the input load goes from low to high, the system exhibits high scalability as the throughput of the system increases from 25,000 to 120,000 events per second, which hints its capability of efficiently processing large volumes of user concurrent activity.

Despite all this added load the system remained low latency with average response times being less than 150 ms despite high traffic. The P95 and P99 latency values indicate a gradual rise but they are within acceptable real-time limits, proving that the architecture can be used for time-critical recommendation and search services. Finally, the resource utilization metrics show efficient use of resources in the infrastructure where CPU, memory, and network utilization scales proportionally with the workload without exceeding a critical threshold. In all, the results validate the proposed

system is robust, scalable and fit for the purpose of the real world, large scale, e-commerce world.

#### D. Discussions

The proposed cloud based architecture shows great potential for real-time intelligent e-commerce search and recommendations. By combining streaming data ingestion, distributed cloud storage, scalable data processing engines, feature Engineering pipeline and service orchestration, the framework can effectively solve the problems associated with handling large scale, high velocity user data. The experimental evaluation shows that the system has high throughput even in the case of heavy workloads, while keeping the response latency within real time limits. This confirms the appropriateness of the architecture to dynamic e-Commerce environments where rapid personalization is critical.

Despite its advantages, the proposed system has a number of limitations. It needs significant cloud infrastructure and computing resources, and this may not be affordable for small or medium enterprises. The amount of complexity of integrating many distributed components makes deployment and maintenance a challenge. Furthermore, the evaluation is carried out based on one benchmark dataset, which may not adequately represent the nature of real-world e-Commerce behavior diversity. Continuous user data collection may also raise privacy, data security and regulatory compliance issues. Future efforts should focus on validation of the system in other domains and lightweight and privacy-preserving solutions.

The proposed framework offers an e-commerce platform scalable solution to improve product discovery and personalization. Real-time data processing enables businesses to respond in real-time to user preferences to improve customer satisfaction and drive up conversion rates. The modular cloud-based architecture allows for flexible scaling at peak time and low latency. Feature stores and APIs make it easier to deploy and update models which reduce the downtime of operations. Moreover, the system supports data-driven decision-making in terms of inventory management, marketing strategies, and customer retention, thus making it very useful for modern-day digital commerce applications.

#### V. CONCLUSION AND FUTURE DIRECTIONS

This study introduces a comprehensive architecture of cloud-based real-time E-Commerce data processing and intelligent recommendations. By the combination of streaming ingestion, distributed cloud storage, scalable processing frameworks, feature engineering pipelines, and containerized orchestration, the proposed system provides a good solution to the challenges related to data volume, velocity, and system reliability. Experimental results show that the architecture achieves a better predictive accuracy and low latency, high scalability, and fault tolerance, which proves the robustness of the architecture in the large-scale commercial environment. The combination of the real-time index, feature store, and the ability to scale down is available in a cloud service, providing cost-efficient resource management, while consistency between the offline learning process and online inference is maintained. These capabilities make the framework a good fit for e-commerce systems with high traffic and requiring fast and personalized user experiences. Reinforcement learning to

adapt recommendations dynamically and explainable AI for greater transparency and privacy-preserving techniques such as Federated learning will be the focus of future research. In addition, the extension of the framework with support for multimodal data such as images, videos and voice queries will further enable personalization and system intelligence.

## References

- [1]. Ojika, Favour Uche, Wilfred Oseremen Owobu, Olumese Anthony Abieba, Oluwafunmilayo Janet Esan, Bright Chibunna Ubamadu, and Andrew Ifesinachi Daraojimba. "Enhancing user interaction through deep learning models: A data-driven approach to improving consumer experience in e-commerce." *Journal of Frontiers in Multidisciplinary Research* 4, no. 1 (2023): 126-137.
- [2]. Liu, Duen-Ren, and Ya-Yueh Shih. "Hybrid approaches to product recommendation based on customer lifetime value and purchase preferences." *Journal of Systems and Software* 77, no. 2 (2005): 181-191.
- [3]. Zhang, Yiman. "The application of e-commerce recommendation system in smart cities based on big data and cloud computing." *Computer Science and Information Systems* 18, no. 4 (2021): 1359-1378.
- [4]. Abdul Hussien, Farah Tawfiq, Abdul Monem S. Rahma, and Hala Bahjat Abdul Wahab. "Recommendation systems for e-commerce systems an overview." In *Journal of Physics: Conference Series*, vol. 1897, no. 1, p. 012024. IOP Publishing, 2021.
- [5]. Ibrahim, Waleed, and Van Kiet Luong. "Transforming E-commerce: A Microservices-Based Taxonomy for Cloud Applications." In *International Conference on Advances in Computing Research*, pp. 476-486. Cham: Springer Nature Switzerland, 2025.
- [6]. Malhotra, Dheeraj, and Omprakash Rishi. "An intelligent approach to design of E-Commerce metasearch and ranking system using next-generation big data analytics." *Journal of King Saud University-Computer and Information Sciences* 33, no. 2 (2021): 183-194.
- [7]. Liu, Liping. "e-Commerce personalized recommendation based on machine learning technology." *Mobile Information Systems* 2022, no. 1 (2022): 1761579.
- [8]. Shankar, Achyut, Pandiaraja Perumal, Murali Subramanian, Naresh Ramu, Deepa Natesan, Vaishali R. Kulkarni, and Thompson Stephan. "An intelligent recommendation system in e-commerce using ensemble learning." *Multimedia Tools and Applications* 83, no. 16 (2024): 48521-48537.
- [9]. Peng, Nianjiao, Xunyong Xiao, Wu Di, Li Ang, and Manhong Wen. "Design and implementation of an intelligent recommendation system for product information on an e-commerce platform based on machine learning." In *International Conference on Internet of Things and Machine Learning (IoTML 2023)*, vol. 12937, pp. 367-375. SPIE, 2023.
- [10]. Tran, Duy Thanh, and Jun-Ho Huh. "New machine learning model based on the time factor for e-commerce recommendation systems." *The Journal of Supercomputing* 79, no. 6 (2023): 6756-6801.
- [11]. Chen, Rongrong. "Research on an intelligent system for optimizing the ranking of e-commerce product search results using deep learning algorithm." In *International Conference on Mechatronics and Intelligent Control (ICMIC 2024)*, vol. 13447, pp. 390-395. SPIE, 2025.
- [12]. Jiang, Liaoliang, Yuting Cheng, Li Yang, Jing Li, Hongyang Yan, and Xiaoqin Wang. "A trust-based collaborative filtering algorithm for E-commerce recommendation system." *Journal of ambient intelligence and humanized computing* 10, no. 8 (2019): 3023-3034.
- [13]. Hanafi, and Burhanuddin Mohd Aboobaidar. "Word Sequential Using Deep LSTM and Matrix Factorization to Handle Rating Sparse Data for E-Commerce Recommender System." *Computational intelligence and neuroscience* 2021, no. 1 (2021): 8751173.
- [14]. Shekhar, Shashi, Anshy Singh, and Avadhesh Kumar Gupta. "A deep neural network (dnn) approach for recommendation systems." In *Advances in Computational Intelligence and Communication Technology: Proceedings of CICT 2021*, pp. 385-396. Singapore: Springer Singapore, 2022.
- [15]. Lutz, Yunus, Timo Wilm, and Philipp Duwe. "Industry Insights from Comparing Deep Learning and GBDT Models for E-Commerce Learning-to-Rank." In *Proceedings of the Nineteenth ACM Conference on Recommender Systems*, pp. 983-986. 2025.
- [16]. Wu, Chunyan, Zemei Liu, and Zhaocui Li. "Identification of Malicious E-commerce Users Based on User Rating Behavior and GNN." *Informatica* 49, no. 10 (2025).