



The Scientific Temper

JULY 2026 Vol. 17 (7): 6540-6556

ISSN: 0976-8653, E-ISSN: 2231-6396

DOI: 10.58414/SCIENTIFICTEMPER.2026.17.7.2493

<https://scientifictemper.com/>

RESEARCH ARTICLE

A Novel Adaptive Dynamic Multi-Objective based Fuzzy Transaction Optimization based effective pattern mining for Customer Basket Analysis

T. JAVITH HUSSAIN¹, K.N. ABDUL KADER NIHAL²

^{1,2} Jamal Mohamed College (Autonomous), Affiliated to Bharathidasan University, Tiruchirappalli – 620020, India.

Corresponding Author*: tjh@jmc.edu

ABSTRACT

In the field of pattern mining (PM), customer basket analysis (CAB) is crucial for understanding buyer behavior and enabling efficient transaction based decision-making. These days, several techniques have been developed for efficient PM, yet they suffer from persistent drawbacks including high computational complexity, significant execution times, and difficulties in handling data sparsity across large-scale item transactions. Therefore, this work proposes a novel Adaptive Dynamic Multi-Objective based Fuzzy Transaction Optimization (ADMO-FTO) technique for enabling efficient CBA based on Dwarf Mongoose Optimization Algorithm (DMO). This work has integrated the adaptive fuzzy weighting mechanism based item occurrence frequency within customer basket. The feedback based reward function based dynamic updation gives addition weightage to balance frequent and rare items under different sparsity levels. The performance of the ADMO-FTO is evaluated based on traditional PM techniques (Apriori and FP-Growth), as well as meta-heuristic based PM (GA-APM, PSO, and ACO). The experimental output shows that the proposed ADMO-FTO outperforms the existing techniques by achieving higher accuracy (96%); lower execution time; improved scalability. At last, the proposed work presents a successful PM method that is especially useful for analysing client purchasing baskets.

Keywords: Pattern Mining, Customer Basket analysis, Genetic Algorithm, PSO, Dwarf Mongoose Optimization, Multi-objective.

INTRODUCTION

The modern dynamic rise of e-commerce and retail systems has generated enormous amounts of consumer purchase data, creating new possibilities for CBA and enabling the extraction of important business insights. By finding connections between purchased items; CBA helps merchants with item recommendations; inventory management; cross-selling strategies. In order to find hidden purchasing linkages inside transaction records, frequent PM and association rule learning have become crucial methods. Recent studies have significantly enhanced market basket analysis through intelligent frameworks that use clustering and deep learning techniques to improve pattern detection and

customer behaviour interpretation (Surana et al., 2026), (Zhang & Ju, 2026).

However, the efficiency of conventional analytical techniques is still threatened by the growing size, diversity, and complexity of contemporary transaction databases. Customer transaction are rarely consistent in real-world retail settings because of different purchasing habits, item occurrences are often distributed unevenly, and many transactions only comprise a small percentage of the available products. Therefore, it is still a big issue to discover large purchasing trends while maintaining efficiency. Several existing techniques, especially frequent itemset mining, focus on optimization strategies to improve pattern extraction (Wu, 2025), (Wan & Han, 2024). In a similar way, decision trees,

association rules, and formal concept analysis have been employed in numerous customer behavior-based studies to uncover temporal trends in decision-making and purchasing (Pitka et al., 2025). However, duplicate patterns and decreased mining accuracy are often the consequence of noisy transactions, overlapping consumer interests, and unclear item linkages.

Therefore, the modern PM techniques are developed on distributed large-scale environment for improving scalability of basket analysis (Sun et al., 2024). Moreover, by integrating utility information into the mining process, high-utility itemset mining techniques have expanded on conventional frequency-based analysis (Yang et al., 2025), and studies on transaction heterogeneity have demonstrated the substantial impact of dataset characteristics on mining success (Trasierras et al., 2024). In addition, recommendation-based methods that combine association rule mining (ARM) and modified RFM analysis have improved customer-centric decision support (Chen & Gunawan, 2023). Thus, the need for more flexible and reliable mining techniques is highlighted by comparative studies that show “frequent PM” algorithms’ effectiveness varies significantly across various transaction contexts (Kallay & Mihoc, 2025). In addition, to enhance retail intelligence and customer engagement, current research has investigated transaction performance characteristics, online shopping behaviour, and tailored recommendation systems (Choi & Park, 2025), (Nakano, 2023), (Chien & Cadeaux, 2026).

These methods have made significant progress, but they mostly concentrate on frequency, utility, scalability, or client segmentation, paying little attention to transaction uncertainty in the process of finding patterns. The capacity to accurately simulate uncertain purchase behaviour prior to pattern extraction becomes crucial for enhancing mining quality and minimising redundant knowledge production as transaction databases get more complicated. Furthermore, advanced utility-based PM techniques still face scalability and efficiency problems when handling large and high-dimensional transactional data (Zhao et al., 2026). This work proposes an Effective PM methodology for CBA based on Fuzzy Transaction Optimization, motivated by these limitations. The proposed approach aims to provide a scalable solution for complex retail transaction settings, enhance the detection of significant item links, boost mining accuracy, and optimise transaction representation through fuzzy modelling. The contribution of the proposed work is shown below.

- Introduced a fuzzy weighting system that dynamically modifies the frequency of item occurrence.
- Reward-based updates are integrated to balance common and uncommon items in different sparsity scenarios. By improving transaction representation, this process obtains fewer redundant patterns.
- An adaptive weighted multi-objective optimization is introduced by concurrently enhancing scalability, runtime, and accuracy.
- Intelligent Swarm intelligence aware role based optimization, where the alpha guides weighted Fuzzy membership updation for item; the scouts investigate potential itemsets by introducing adaptive variety; babysitters ensure robustness by maintaining stability during multi-objective mechanism

LITERATURE REVIEW

Huang et al. (2021) presented a distributed technique for quickly identifying recurring patterns in large amounts of data in 2021. Their distributed architecture significantly reduced runtime, allowing them to execute more than 40% faster than Apriori. Although performance decreased in noisy and diverse contexts, this work's strength is its scalability over large datasets. This work highlighted the significance of resilience in various contexts and demonstrated how parallelization might expedite frequent PM.

In 2023, the Profit-Support Fuzzy ARM (P-FARM) Dogan (2023) was introduced by Dogan, who claimed that his contribution increased accuracy by 12% compared to the tradition FARM method. However, the approach is particularly limited to profit-based analysis and does not ensure adaptability across different datasets. Therefore, more improvements are required to enhance e-commerce customization by ensuring greater adaptability.

Smaili and Hachimi (2023) introduced the RFM-D classification model in 2023 to enhance customer response prediction. The outcome of their study obtained 92% of accuracy in customer segmentation and response analysis, whereas existing methods achieved only 85% of accuracy. However, scalability still a significant challenge when handling large transaction datasets. Therefore, improving the efficiency of classification models became a primary need.

Conversely, Suchacka and Chodak (2017) developed a novel association rules that focused on customers purchase behaviour analysis based on online purchase. Although their model achieves 88% of accuracy, the considerable challenges of the work was handling contemporary large-scale e-commerce datasets.

However, the proposed worked achieved high probability-based purchase modelling still it is limited to poor generalization.

In 2023, Zhu et al. (2023) focused on obtaining correlation knowledge for distribution network. Their strategy demonstrated correlation mining's usefulness outside of retail by increasing planning efficiency by 15%. The results demonstrated how data mining can enhance energy system decision-making, despite its restricted relevance to basket analysis. The contribution of the proposed work highlights the cross-domain applicability of PM but limited to shopping cart analysis.

Similarly, in the PM research the authors Fouad et al. (2021) have presented a dependable high-utility PM framework in 2021. In their research they achieved 94% of accuracy while the existing technique scored 89% with HUP-Growth. But the computation overhead and susceptibility to parameter adjustment are the challenge while working large dataset.

Chen and Gunawan (2023) introduced an efficient RFM analysis with association rules to improve store recommendations. In their hybrid approach, the accuracy is highlighted as 10% higher than the conventional RFM method. This study provided useful enhancements for retail recommendation systems by demonstrating the advantages of combining RFM and rule mining. However, its performance is not proven as an efficient model outside of the retail application.

On the other hand, Karulkar et al. (2025) have presented a hybrid approach based on fuzzy logic and K-Means cluster for customer PM. As other previous methods, their approach achieved 91% of accuracy while demonstrating a hybrid approach. This scenario's demonstrates that the sensitivity to clustering effects reduced resilience due to sparse data. Moreover, the proposed shows insights on hybrid clustering-fuzzy algorithms for consumer analytics and is fairly thorough.

In the same year, Agarwal et al. (2025) have introduced a multi-phase recommender architecture for pattern warehousing based on evolutionary optimization that achieves 95% of accuracy with minimal redundancy. Moreover, the multi-phase architecture ensures balanced accuracy and computational efficiency. In support to PM, Dogan et al. (2022) introduced a fuzzy ARM in 2022 for e-commerce that achieves 90% accuracy.

Although current methods for extracting frequent and utility-based patterns improve recommendation accuracy, they exhibit shortcomings regarding robustness, scalability, and heterogeneity when applied to large transactional datasets. Therefore, the challenge lies in developing an adaptive framework that strikes a balance between identifying both frequent and rare items, while simultaneously ensuring the effectiveness and reliability of market basket analysis.

Table 1: Tabulation of Literature Studies

S.No	Author(s)	Proposed Technique	Limitation
1.	Huang et al. (2021)	Distributed method for fast mining frequent patterns from big data	High variability limits scalability, and noisy data clusters lead to a loss of efficiency.
2.	Dogan (2023)	P-FARM	Focused mainly on profit-support; limited generalization across diverse datasets
3.	Smaili & Hachimi (2023)	New RFM-D classification model	Limited to predicting customer responses; lacks scalability with large datasets.
4.	Suchacka & Chodak (2017)	Association rules for purchase probability assessment	The scope of the dataset is outdated; adaptability to modern, large-scale e-commerce is limited.
5.	Zhu et al. (2023)	Correlation knowledge extraction for distribution network planning	Domain-specific (computational power); direct application is limited to shopping cart analysis.
6.	Fouad et al. (2021)	Reliable High Utility PM	High computational cost for very large datasets; sensitive to parameter changes.
7.	Chen & Gunawan (2023)	Modified RFM + Association Rules	Limited to retail transaction datasets; not transferable to other domains.
8.	Karulkar et al. (2025)	K-Means + Fuzzy Logic for consumer pattern analysis	Sensitive to clustering parameters; limited robustness with distributed data.
9.	Agarwal et al. (2025)	Multi-Phase Recommender Framework using Evolutionary Optimization	Complexity increases with dataset size; requires fine-tuning of evolutionary parameters.
10.	Dogan et al. (2022)	Fuzzy ARM for e-commerce product association	Focused on revenue; limited consideration of usability aspects and distributed data.

PROPOSED METHODOLOGY

In the proposed work, a novel ADMO-FTO is introduced for enabling a high-level pattern mining. This framework effectively manage several challenges (sparse, imbalanced, and heterogeneous transaction datasets) that affect the e-commerce environments. Several traditional techniques including Apriori and FP-Growth that rely on rigid binary encodings and exhaustive rule generation. These technique often fail to capture weak correlations and struggle with scalability. Therefore, ADMO-FTO based fuzzy integrated Adaptive Dynamic Multi-Objective Optimization guided Dwarf Mongoose Optimization (DMO) algorithm is developed.

In this framework, an adaptive weighted fuzzy membership function is introduced for item co-occurrence frequency, ensuring that both frequent and rare items are represented fairly. After that, these fuzzy weights are iteratively improved by ADMO that uses swarm intelligence from DMO for dynamic parameter adaptation while balancing several objectives, including accuracy, sparsity handling, and processing efficiency. Moreover, the proposed work introduces an IOA-based frequency modelling approach for quantifying item occurrence in transactional data. It incorporates sparse matrix analysis to capture structural data imbalance and facilitate adaptive representation learning.

The overall flow of the proposed work is given in Figure 1. In addition, a fuzzy-aware transaction model is developed by integrating normalized IOA frequencies into binary encoding to enhance pattern mining mechanism.

PROBLEM FORMULATION

The primary challenges associated with online sales and grocery transaction datasets include (i) high dimensionality, (ii) data sparsity, and (iii) class imbalance in item occurrence distributions. These features highly restrict the efficiency of traditional binary and frequency-based representation models due to they are inefficient to handle structural variation and uncertainty in item relevance. The primary goal is to develop an adaptive representation of objects that captures their contextual value while maintaining transactional structure under sparse settings. Formally, let the transactional database be described as $(D = \{T_1, T_2, T_3, \dots, T_N\})$, where each transaction $(T_i \subseteq I)$ is taken from the universal itemset $I = \{i_1, i_2, i_3, \dots, i_N\}$. However, three primary challenges are arising while constructing the transaction model such as sparsity and high dimensionality, Imbalanced Item Occurrence Distribution, and Sensitivity of Parameters in Representation Learning.

Sparse and high-dimensional transactional data significantly degrade the performance of conventional mining and representation techniques. Parameter sensitivity and imbalance between dense and sparse regions further exacerbate learning instability and hinder accurate pattern extraction. The performance of traditional mining and representation approaches is severely affected by sparse and high-dimensional transactional data. And the Learning instability is increased and precise pattern extraction is limited by parameter sensitivity and an imbalance between dense and sparse regions.

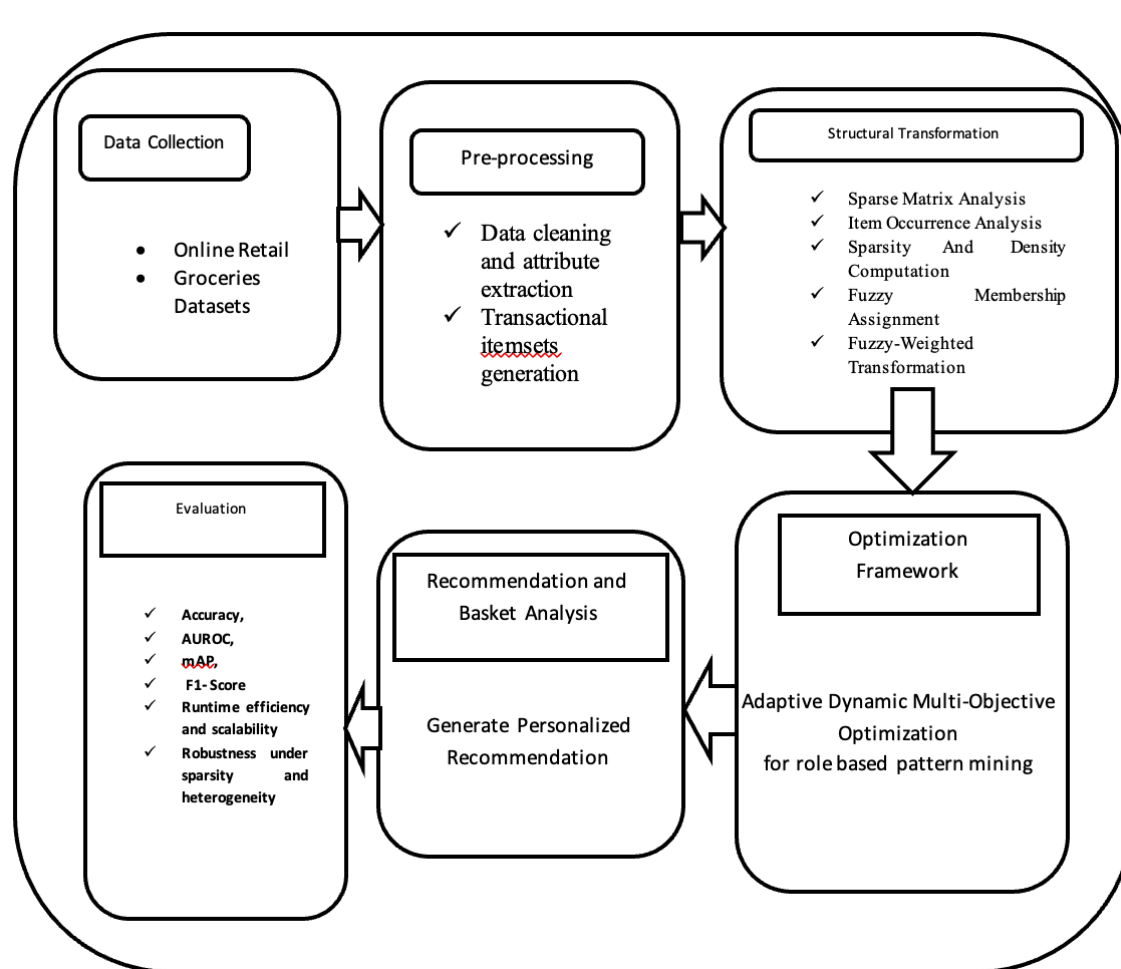


Figure 1: Overall Flow of the Proposed Work

PRE-PROCESSING

The pre-processing steps performs a series of operations to obtain structured customers' baskets representation for adaptive frequent PM. In the pre-processing steps, Data cleaning, Attribute extraction, Transactional Itemset Generation, Encoding, Item Occurrence Analysis, Sparse Matrix Analysis, Fuzzy-weighted Transformation are mainly considered for transforming the datasets into suitable format for further processing.

Data Cleaning and Attribute Extraction

In the initial data cleaning step, data inconsistency and noisy transactional records properly are removed for ensuring the reliability and semantic consistency. To prevent inadequate item association, the removal of missing product descriptions and cancelled items are primarily executed at the time of pattern extraction. In addition, eliminating unnecessary symbols ($\times$, $+$, δ , $"$, $\backslash$ and $;$) are executed for normalizing the product description for improving the textual consistency during transactional analysis. Moreover, the repeated and irrelevant terms are primarily focused and removed from the transactional records to maintain semantic

record. Here, when comparing with online Retail dataset (OR-DS), Groceries Dataset (GR-DS) maintains inherently organized transactional record that reduces fewer refinement operations. In this step, the essential attributes are extracted from the dataset to enable CBA. In this case, the role of invoice and product description are the significant elements that directly influence in customer behaviour; therefore, they must be considered within transactions. Moreover, these characteristics enables effective transaction modelling while eliminating unnecessary attribute from the transaction record to obtain effective patter mining activates.

Itemset Generation and Encoding

As discussed, the transaction item sets (customer basket) are generated based on the products that are directly associated with invoice. Therefore, each invoice within a transaction item sets maintains distinct customer transactions, including multiple product purchase details without conflict. Conversely, GR-DS already maintains predefined customer transaction that avoids the steps required in OR-DS. Thus, the transactions generated in this way provide the necessary input data to analyze the customer's

basket and extract frequent patterns. The raw generated transactional itemsets are unsuitable for direct computation operations that requires binary transaction item matrices (numerical transformation) using one-hot encoding. Here, distinct products within the dataset are represented by columns, whereas customer transactions are represented by rows. In customer transactions, a binary value is assigned based on the presence or absence of a product. The resulting binary transaction matrix provides a robust structural representation for supporting evaluation, transaction analysis, and model extraction for optimization. The following equation (1) shows the mathematical formulation of the binary-transaction-item matrix.

$$T = [t_{ij}]_{m \times n} \quad (1)$$

Where "m" is the number of transactions, "n" is the total number of unique items, and " $t_{ij} \in [0,1]$ " specifies whether item j is present in transaction "i" or not. This shows that in the transaction-item matrix, the, $t_{ij} = 1$ signifies item presence and the, $t_{ij} = 0$ indicates item absence.

Sparse Matrix Analysis

The purpose of sparse matrix analysis is to describe the density distribution of transactions in customer baskets by utilizing binary representations of transactions and items. In this case, the majority of transactions only cover a small portion of the catalogue's items that produce extreme sparse matrices. This stage mainly covers Item Occurrence Analysis and Sparsity and Density Metrics.

Item Occurrence Analysis

Item Occurrence Analysis (IOA) is an important stage in the pre-processing phase to execute support-based statistical evaluation for frequency extraction. In fact, the IOA is referred to as called support value computation shows how frequently a particular item occurs overall customer transactions is shown in equation (2):

$$IOA(X) = \frac{count(X)}{m} \quad (2)$$

In this case, X stands for item, m for transactional records ($D = \{T_1, T_2, T_3, \dots, T_m\}$), where T_i for an ith transaction, and $count(X)$ represents the number of transactions that contain the itemset X. Consider ($I = \{i_1, i_2, i_3, \dots, i_n\}$) illustrates the set of all items. An itemset is denoted as $X \subseteq I$, denoting a subset of items drawn from this set. The

estimated values offer statistical insight into dominant transactional patterns within the datasets and identifies high-frequency purchase items.

Sparsity and Density Metrics

The distribution of non-zero items in the transaction item-matrix is typically represented quantitatively at this stage. This phase mainly contribute in generating Adaptive fuzzy weighting strategies to capture the differences between sparse and dense ratio.

Sparse Ratio

The mathematical computation of the sparse ratio of the transaction matrix is computed using equation (3).

$$S_T = 1 - D_T \quad (3)$$

In fuzzy representation learning, sparsity measures are employed to describe dataset structure and direct adaptive weighting sensitivity. The transaction-item matrix's sparsity ratio is represented by S_T . The number of non-zero entries is denoted by $NNZ(T)$ and the "m" represents the total number of customer transactions. Meanwhile, the total number of unique items is represented by "n". Finally, the total number of matrix elements is represented by $m \times n$. The matrix's non-zero entries are calculated using equation (4).

$$NNZ(T) = \sum_{i=1}^m \sum_{j=1}^n t_{ij} \quad (4)$$

Density Measure

The mathematical formulation of transaction density matrix is shown in equation (5)

$$D_T = \frac{NNZ(T)}{m \times n} \quad (5)$$

The density of the transaction density matrix is defined as D_T , representing the ratio of non-zero item occurrences to all possible transaction item combinations. A lower density means that the majority of the matrix's entries are zero-valued, suggesting that consumer usually only buy a small portion of the total product space. Meanwhile, the transaction matrix has a high sparsity ratio, indicating that most of its entries are "0". As a result, it shows that only a small subset of items are involved in customer's purchasing behaviour. Therefore, the proposed work performs the sparsity behaviours analysis technique before processing the transaction modelling operation.

Table 2: Illustration of Transaction Matrices

Dataset	Transactions (m)	Items (n)	Non-zero entries $NNZ(T)$	Sparsity Ratio S_T	D_T
Online Retail	25,900	4,223	259,000	0.9976	0.0024

Groceries	9,835	171	43,367	0.9739	0.0261
-----------	-------	-----	--------	--------	--------

Table 2 shows how sparse the transaction–item matrices are in both datasets. For the OR-DS, the number of non-zero entries $NNZ(T)$ is about 259,000 that translates to a S_T of 0.9976 and a density of D_T 0.24%. The GR-DS, with 43,367 $NNZ(T)$, is slightly denser at 0.26%, but still highly sparse with a ratio of 0.9739. Together, these results confirm that both datasets contain very few co-occurring items, reinforcing the need for fuzzy weighting and adaptive modelling to uncover meaningful purchasing patterns.

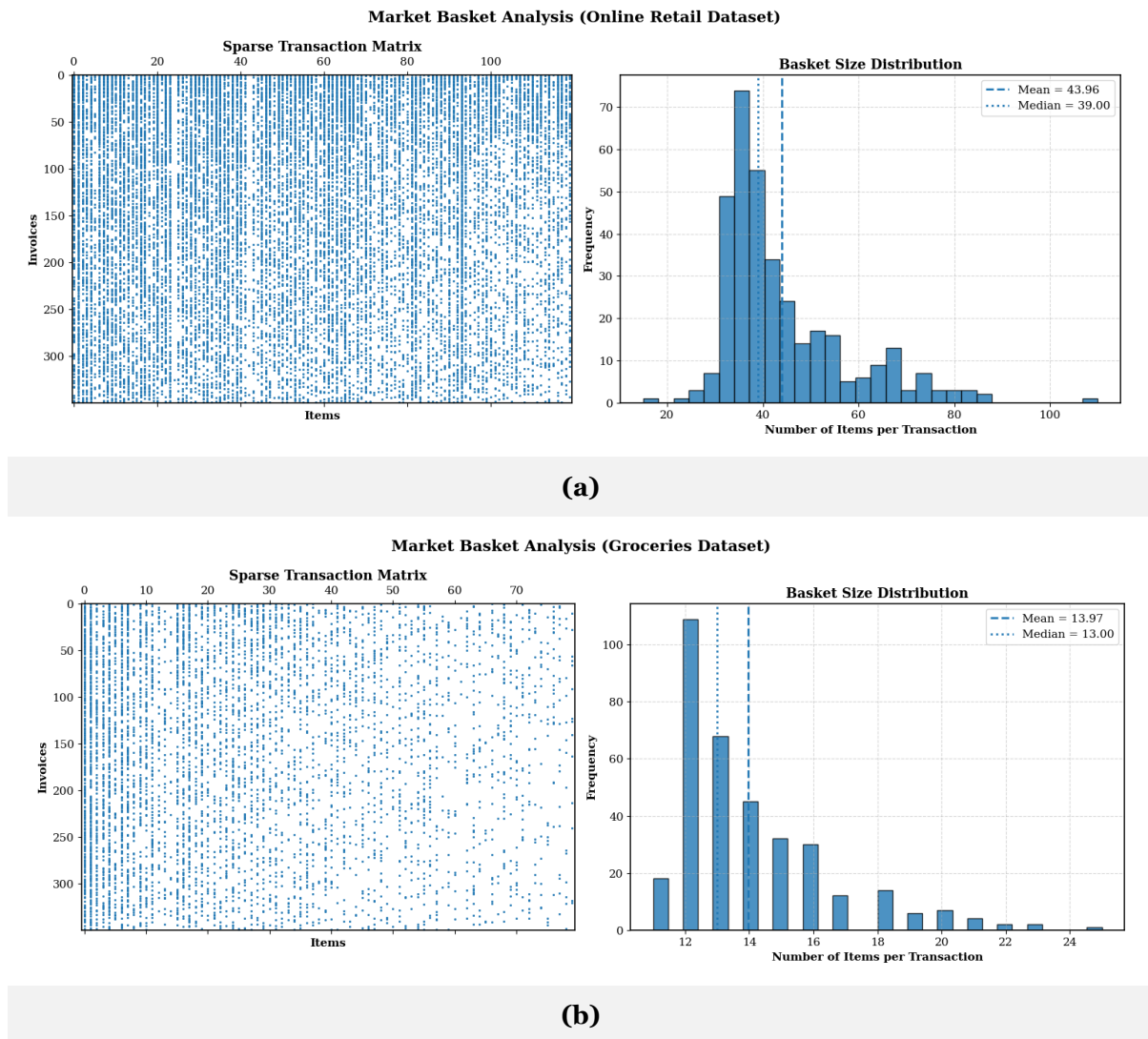


Figure 2: illustrates the graphical visualization of market Basket analysis

Figure 2(a) visualizes the basket size (BS) distribution within the OR-DS. It illustrates how the majority of transactions contain only a few items that typically fewer than 10 items and how the distribution, characterized by a long tail, occasionally includes large carts containing more than 20 items. Because of the OR-DS's extremely sparse transaction-item matrix, it is challenging to find consistent correlations across diverse baskets. The grocery dataset is shown in Figure 2(b) in which BSs are more concentrated: the

majority of transactions include three to five items, with very few exceeding ten. This comparison shows that supermarket data has denser and more stable transaction patterns, while internet retail data is characterised by sparsity and variability. This justifies the use of adaptive fuzzy weighting techniques to robustly model purchasing behavior across diverse contexts.

FUZZY WEIGHTING MODEL

Membership Function Design

The traditional binary transactional indicates whether an item is present ("1") or absence ("0") within the baskets but it is not efficient in capturing the relative transactional significance of individual products. To overcome this limitation, a fuzzy membership function is introduced for item j , which is mathematically defined in equation (6).

$$\mu_j = \frac{O_j}{\max(O)} \quad (6)$$

Where O_j denotes the occurrence frequency of item j . As discussed, estimating the item significance via the IOA values that is denoted by O_j , and its fuzzy membership value is represented as $\mu_j \in [0,1]$. μ_j is represented as the global fuzzy importance of item j . By adaptively prioritising informative items, this method reduces the impact of transactional patterns that are less important. Membership function can be reliably incorporated into transaction modelling by normalization that ensures stability and compatibility.

Transaction Weight Assignment

The fuzzy representation learning strategy is adopted into the transactional modelling by assigning fuzzy weights to transactional items based on their occurrence frequency. The transactional frequency of individual products across customer baskets is initially performed using item occurrence analysis. In this process, higher occurrence frequency items are considered to be more informative for the analysis of purchasing behaviour and the identification of adaptive patterns. The equation (7) highlights the mathematical formulation of fuzzy-weighted transaction representation.

$$f_{ij}^w = T_{ij} \times \mu_j \quad (7)$$

Where f_{ij}^w illustrates the fuzzy-weighted transaction value, T_{ij} represents the entry of the binary transaction-item matrix. Moreover, this equation is based on the baseline fuzzy-weighted transaction model where the membership's values are considered static while executing learning process. As a results, the transformed fuzzy-weighted representation enhances the capability of the proposed model to identify dominant purchasing relationships from sparse customer basket datasets. However, the fuzzy-weighted representation is still susceptible to sparsity and parameter tweaking despite these advancements that is the reason for developing an adaptive optimization framework is incorporated.

Optimization Framework

As discussed in previous section, fuzzy weighting enables significant enhancement in transactional representation and offers appropriate decision making. However, its performance is still limited by sparsity and parameter sensitivity. Therefore, there is a need to develop and adaptive optimization algorithm that enables robust pattern discovery while ensuring seamless integration between dynamic multi-objective strategies and fuzzy transaction learning (FTL).

Adaptive Dynamic Multi-Objective Strategy

The proposed work introduces an adaptive dynamic multi-objective optimization (ADMO) strategy to address the issues associated with sparse and high-dimensional transactional data in which static parameter settings generally fail to produce optimal results. This section adopts the ADMO strategy based on the Adaptive Dwarf Mongoose Optimization (DMO) algorithm, ensuring that the fuzzy transaction model benefits from dynamic multi-objective balancing and swarm intelligence. Moreover, the multi-objectives problems like accuracy, sparsity, computational stability, and scalability are mainly considered and addressed via adaptive mechanism using multi-objective reward function (is shown in equation (8)).

$$R = \sum_{k=1}^K w_k \cdot f_k \quad (8)$$

The f_k in equation (8) expresses the multi-objective functions such as accuracy (f_1), minimize sparsity (f_2), minimize instability (f_3), and maximize scalability (f_4). These function dynamically updated via adaptive weights function (9) based on the performance feedback.

$$w_k^{t+1} = w_k^t + \eta \cdot \Delta_k \quad (9)$$

In equation (9), the adjustment controller (or performance feedback controller) is represented as Δ_k and learning rate is represented as (η) and Δ_k . With this adaptive mechanism, the multiple objectives problems are balanced and maintained in a way similar to a single-objective problem. The dynamic weight adjustment mechanism achieves robust convergence and enhances resilience against noisy transactional patterns. Moreover, FTL improves the overall performance by effectively integrating with the dynamic weight adjustment mechanism.

Fuzzy Transaction Learning Integration

As discussed, the primary issues such as sparsity and parameter sensitivity are limits the efficiency of traditional fuzzy-weighted representations in handling transactional data. Therefore, an adaptive dynamic multi-objective (ADMO) strategy is developed to optimally update the fuzzy

membership values μ_j , ensuring that informative items are reinforced while less significant patterns are suppressed. The mathematical formulation of this process is presented in equation (10).

$$f_{ij}^{w^{opt}} = T_{ij}^{opt} \times \mu_j^{opt} \quad (10)$$

Similar to equation (7), equation (10) integrates the same parameter but with additional superscript "opt" represent their optimized forms, such as the optimized fuzzy weighted transaction value, the optimized transaction item matrix, and the updated membership weight guided by ADMO's multi-objective balancing mechanism. In addition to capturing item importance, this integration guarantees that the fuzzy transaction model can instantly adjust to sparsity, stability, and scalability needs. The framework improves its capacity to discover dominating purchase associations and achieves robust convergence by directly integrating optimization into the fuzzy learning process.

ADMO-based Fuzzy Transactional Learning Framework

The fuzzy-weighted transaction representation offers a flexible mechanism for capturing item-level relevance from transactional records. However, it faces parameter-related issues (sensitivity and sparsity) under uncertainty and high dimensional environment. Weighted fuzzy-based decision making is enhanced by integrating an optimization technique to improve the fuzzy membership parameters. Still, the inherent data imbalance and high sparsity characteristics in real-world transactional datasets make it difficult to ensure robustness within the overall transaction modelling framework. Therefore, the ADMO guided Adaptive Dwarf Mongoose Optimization algorithm, is employed to optimize the fuzzy membership values based on dataset characteristics ($IOA(X)$, S_T , and D_T). These characteristics are considered for enabling optimal learning mechanism under sparse condition. Particularly, this hybrid approach ensure robust PM by handling high-dimensional and sparse transactional environments. In the following, the background study about the Dwarf mongoose algorithm is discussed.

Background: Dwarf mongoose algorithm

Dwarf Mongoose Optimization (DMO) is inspired by the (social and migration) lifestyle of dwarf mongoose populations. This algorithm performs well for single objective problem but is limited by poor adaptability due to static weighting. To overcome this limitation, a dynamic multi-objective with adaptive weight adjustment mechanism is introduced for ensuring a balanced exploration and exploitation phase. In the DMO algorithm, the populations is categorized into

three groups (Alpha, Scout, and Babysitter) where each group plays different roles. Within the ADMO-FTL framework, the alpha guides optimal fuzzy membership updates, scouts provides adaptive variability for dynamic datasets, and the role of babysitters is maintaining stability for ensuring robust multi-objective optimization in fuzzy transaction. The equation (11) shows how the initialization of the alpha group ($\alpha\delta$) is represented.

$$x_{i,j} = rand_{dist}^{uni}(LB, UB, Dim) \quad (11)$$

The generated initial position is represented by $x_{i,j}$, LB and UB are respectively denoted as lower and upper boundaries. The number of transactional feature in the item universe ($I = \{i_1, i_2, \dots, i_n\}$) is expressed by the dimension (Dim). The candidate itemsets are defined as subsets of this entire sets of items in the transactional database ($X \subseteq I$). The terms " r " and " r_{dist}^{uni} " represent uniform distributed random number. In addition, the female leader of the dwarf mongoose population appears in the ($\alpha\delta$), as shown by equation(12).

$$a = \frac{fit_i}{\sum_{i=1}^N fit_i} \quad (12)$$

The population size is represented by "N" and the fitness value of the each agent(fit_i). In the ($\alpha\delta$), the number of agents is given by $n = N - bs$ ($bs \rightarrow$ the group size of babysitter). The updation rule of Alpha leader is expressed in equation (13) that exploits high-utility candidate itemsets by guiding agents toward promising regions of the search space.

$$X_{new} = X_k + phi \times peep \times (X_i - X_k) \quad (13)$$

The new position of the dwarf mongoose is denoted by X_{new} , the vocalisation coefficient is set to 2 in this study, phi is a random number uniformly distributed between 0 and 1, the current position of the female leader is denoted by X_i , and the position of a random individual in the $\alpha\delta$ that differs from the female leader is represented by X_k . The new position, X_{new} , is then evaluated for fitness, and after obtaining, fit_{i+1} , the mound value is calculated using equation (14).

$$sm_i = \frac{fit_{i+1} - fit_i}{\max\{|fit_{i+1}, fit_i|\}} \quad (14)$$

Where the mound value is represented by sm_i that determines the average mound value is shown in equation (14). The Scout group($\sigma\delta$)'s movement formula is represented by equation (15), and their main duty is to find new sleeping mound locations.

$$X_{i+1} = \begin{cases} x_i - CF \times phi \times r \times [x_i - \bar{M}], & \text{if } \varphi_{i+1} > \varphi_i \\ x_i + CF \times phi \times r \times [x_i - \bar{M}], & \text{else} \end{cases} \quad (15)$$

Where r represents the generation of random numbers between 0 and 1. For the dwarf mongoose population, CF is the mobility parameter, and $\bar{M}$ is the directional vector that determines the direction of the mongoose's movement. Equations (16) and (17) show the calculations for CF and M , respectively.

$$CF = \left(1 - \frac{t}{T}\right)^{\left(\frac{2t}{T}\right)} \quad (16)$$

$$\bar{M} = \sum_{i=1}^N \frac{x_i \times sm_i}{x_i} \quad (17)$$

Where " t " is the current iteration number and " T " is the maximum number of iterations, the movement parameter " CF " decreases linearly with iterations. This incremental reduction guarantees controlled investigation of prospective itemsets in the context of PM. The $\sigma\delta$ investigates novel potential itemsets, while the $\alpha\delta$ concentrates on taking advantage of high-utility itemsets inside the population. When the timing parameter fulfils ($C \geq C_{exchange}$, the Babysitter group ($\beta\delta$) starts a role-based exchange. The $\sigma\delta$ switches positions with the $\alpha\delta$ while Scouts continue to diversify candidate itemsets since the $\alpha\delta$'s exploitation ability is deemed weak under these circumstances. By preserving a balance between exploration and exploitation in fuzzy utility PM, this dynamic exchange enhances the ADMO-FTL framework.

The following PM is based on an optimized matrix of uncertain transactions, generated through the adaptive Dwarf Mongoose optimization algorithm. To identify dominant and rare associations, high-utility frequent itemsets and association rules are generated from this matrix, along with their corresponding rules. In the following step, recommendation and basket analysis is executed for leveraging these mined patterns to transform technical results into relevant customer insights. This dynamic exchange mechanism thus enables the ADMO-FTL platform to execute reliable PM of uncertain value.

Recommendation and Basket Analysis

In consumer basket analysis, the optimised fuzzy transaction matrix produced by the suggested ADMO-FTO architecture provides the basis for practical insights. The system finds both dominating and rare product linkages that traditional approaches frequently miss by mining high-utility frequent itemsets and association rules. Personalized product recommendations,

focused promotions, and enhanced cross-selling opportunities are made possible by the conversion of these mined patterns into recommendation tactics. In addition, the approach facilitates basket analysis by revealing hidden relationships and purchase patterns across diverse datasets. This improves customer engagement and decision assistance in real-world retail settings by ensuring that recommendations are both accurate and robust under sparsity. The following pseudocode give the clear flow about the proposed PM approach.

Pseudocode: ADMO-FTL framework based robust PM

Input

- Transactional Database $D = \{T_1, T_2, \dots, T_m\}$
- Item Universe $I = \{i_1, i_2, \dots, i_n\}$
- Population Size N
- Maximum Iteration T

Output

- Optimized Fuzzy-weighted transaction matrix T^{opt}
- Dominant Itemsets F

Step 1: Pre-processing

- Generate binary transaction-item matrix using equation (1)
- IOA computation using equation (2)
- Sparsity Density Matrix equation (3) to (5)
- Identify candidate itemsets $X \subseteq I$ for mining

Step 2: Fuzzy weighting Model

- Estimate the Item Significance
- Fuzzy membership function computation using equation (6)
- Fuzzy weights assigning using equation (7)

Step 3: ADMO construction

- Define reward function using equation (8)
- Dynamic weights updation using equation (9)
- Optimize fuzzy representation using equation (10)

Step 4: ADMO-FTL based DMO for Robust PM

1. Initialize $\alpha\delta$ position (equation.11).
2. Allocate candidate itemsets for exploitation.
3. Compute normalised fitness (equation.12).
4. Evaluate utility of elements to balance support an reliability.
5. Update alpha position (equation.13).

6. Refine high utility itemsets by exploiting strong associations.
7. Calculate mound value (equation.14).
8. Estimate scout movement (equation.15).
9. Update mobility (equation.16).
10. Compute direction vector (equation.17).
11. Apply babysitter exchange threshold ($C \geq C_h$) to swap alpha nd babysitter roles, stabilizing mining and avoiding premature convergence.

Output

- Return optimized fuzzy transaction matrix T.
- Extract dominant itemsets F.
- Provide personalized itemsets recommendations.

RESULTS AND DISCUSSION

Dataset description

The OR-DS (University of California, Irvine (n.d.)) is based on real-transaction dataset

collected from a UK-based online retail company from 2010 to 2011 (December). This dataset is primarily based on gift products sales, particularly integrating wholesale customers. In facts, the dataset holds 5, 41,909 records that consists of invoice number, product description, stock code, unit price, invoice date, customer id, quantity, and country. As a result, this dataset is used in a multiple fields, such as customer segmentation research, recommendation systems, ARM, and CBA (Chen et al., 2012). On the other hand, the groceries transactional dataset (Hahsler et al., n.d.) found in the arules package is the source of the GR-DS. The datasets integrates 9835 customer sales transactions records for 169 categories of items based on regular purchase details of supermarket. The dataset's transactional structure makes it highly suitable for applying ARM, market basket analysis, and frequent pattern recognition (Hahsler et al., 2006). The statistical summary of the dataset is shown in Table 3.

Table 3: Statistical Dataset Description

Name	Transactions	Items	Types
Online Retail	25,900	4,223	Invoice-based retail
Groceries	9,835	171	Transactional basket

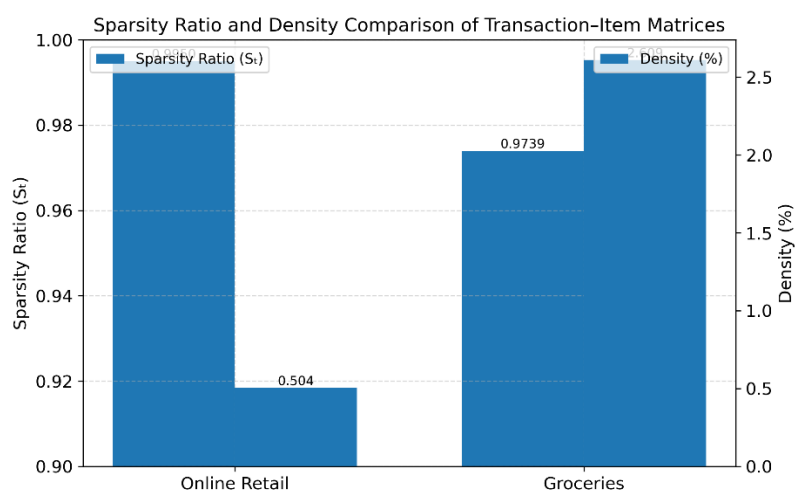


Figure 3: Illustrates the sparsity and density comparison

In Figure 3. two datasets such as Online retail and GR-DSs are compared in terms of density ratio and sparsity. It demonstrates that the OR-DS achieves good sparsity (0.995) and density (0.5 percent). Conversely, the GR-DS has moderate

density (2.6 percent) and sparse (0.974). This highlights the difficulty of identifying significant trends in a sparse transactional context is shown by these complementary measures.

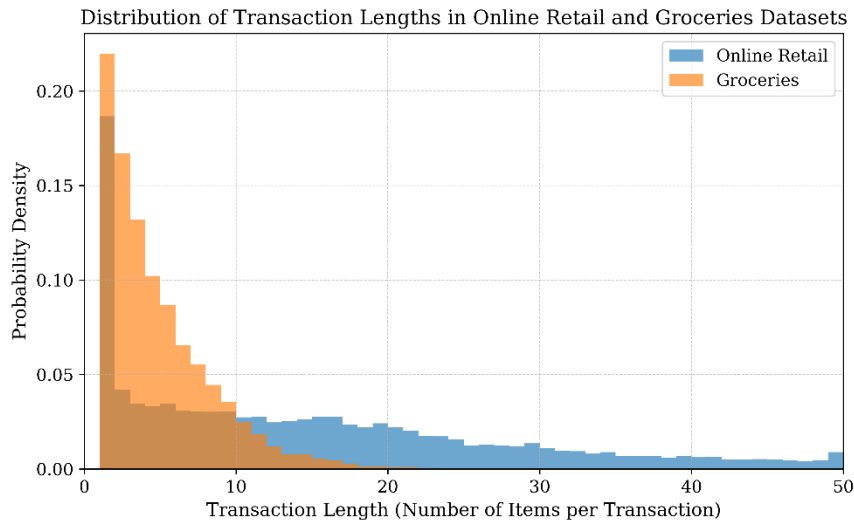


Figure 4: Illustration of Transaction length distribution

The distribution of transaction length in the "Grocery dataset" and "Online retail" datasets is illustrated in Figure 4. In the "Grocery dataset", most transactions involve only a few items, resulting in a greater proportion of small BSs. In

contrast, the online commerce dataset exhibits a wide range and a wider distribution, indicating the sporadic occurrence of independent transactions involving a large number of items.

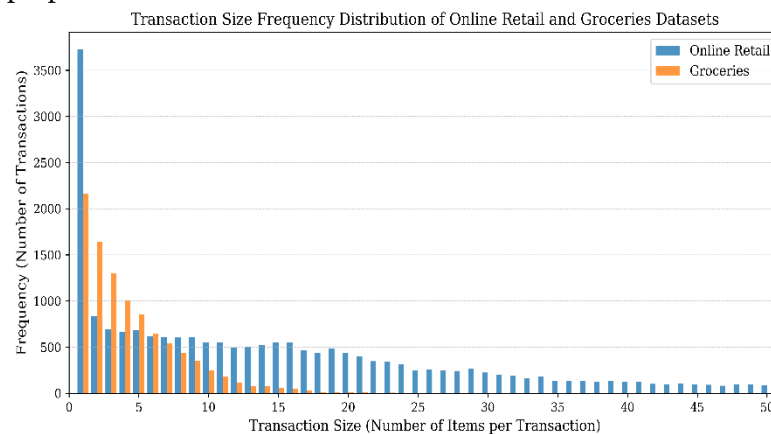


Figure 5: Graphical illustration of Transaction frequency size distribution

Figure 5. Transaction size frequency distribution of Online Retail and GR-DSs. Both datasets exhibit skewness, with the majority of transactions containing only a few items. The GR-DS shows a sharper concentration of short

baskets, while Online Retail demonstrates a broader spread with a longer tail, reflecting occasional transactions with larger numbers of items.

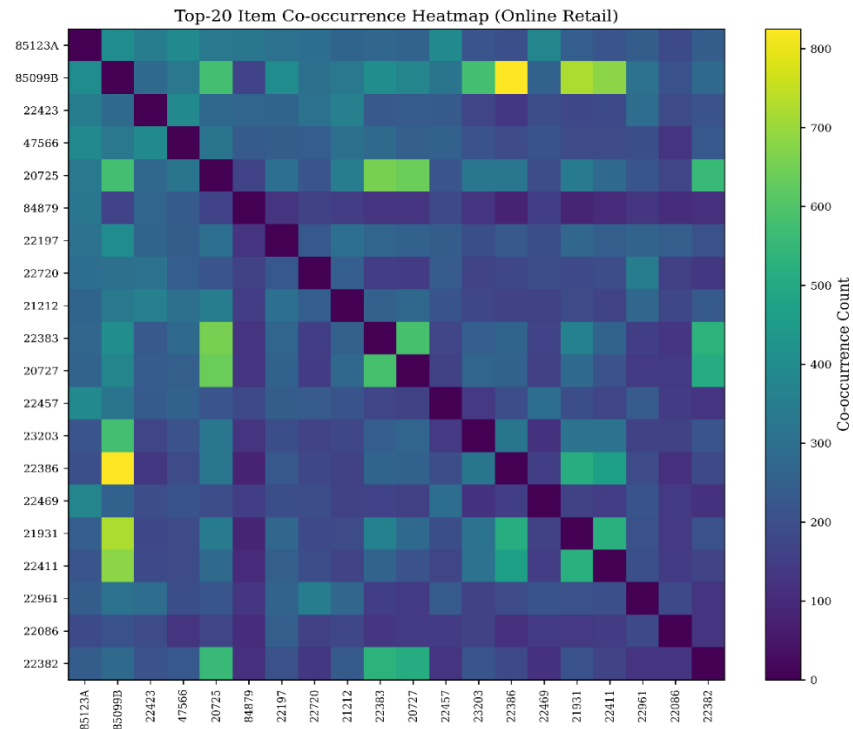


Figure 6: illustrates the item co-occurrence heatmap (OR-DS)

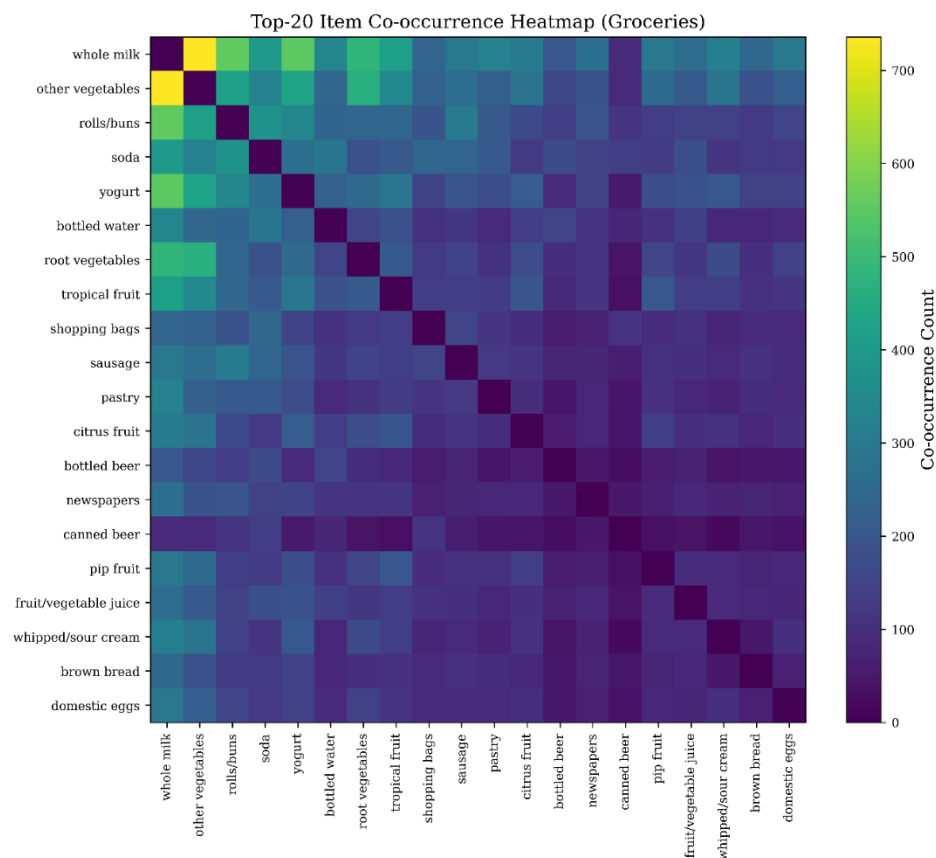


Figure 7: illustrates the item co-occurrence heatmap (grocery dataset)

The proposed methodology is discussed in Section III are directly validated by the item co-occurrence heatmaps based on the OR-DS and GR-DS. In Figure 6 and Figure 7, item

correlations patterns are clearly visualized based on OR-DS and GR-DSs. The diffuse co-purchase pattern with weaker item correlations is shown in Figure 6. Meanwhile, strong co-purchase clusters

emerge among staple items in Figure 7. The proposed approach captures item co-occurrence and methodically captures weak correlations to improve suggestion accuracy.

COMPARATIVE ANALYSIS

The proposed ADMO-FTO framework is evaluated on two primary datasets (Online Retail and GR-DS) using standard evaluation metrics (accuracy, recall, F1-score, Mean Average

Precision (MAP), and Area Under Curve (AUC). The experimental results are summarized in a Table IV and presented in Figure 10. The proposed ADMO-FTO framework is compared against traditional metaheuristic methods, such as GA-APM, PSO, and ACO. This comparison is conducted specifically under conditions of highly sparse data. Under such conditions, traditional algorithms typically yield significantly poorer results, whereas the proposed method demonstrates better performance.

Table 4: Dataset-Wise Comparison of various techniques

Dataset	Metric	GA-APM	PSO	ACO	ADMO-FTO
Online Retail	Precision	92.68	91.76	91.14	96.26
	Recall	89.57	88.66	89.43	94.95
	F1-Score	91.1	90.18	90.28	95.60
	MAP	88.63	90.61	89.11	92.86
	AUC	94.52	93.42	92.42	95.45
Grocery dataset	Precision	92.14	90.66	91.01	96.64
	Recall	88.68	89.97	90.12	94.19
	F1-Score	90.38	90.31	90.56	95.40
	MAP	89.68	89.97	90.44	91.49
	AUC	93.68	92.97	92.11	96.12

As shown in Table 4, ADMO-FTO demonstrates superior precision, recall, F1-score, mAP, and AUC across both datasets. The AUC values obtained for ADMO-FTO range between 95% and 96%, significantly outperforming other existing methods. Most existing methods exhibit

performance levels below 90%. The AUC is 96.12 on the "Grocery" dataset and 95.45 on the "Online Retail" dataset. Furthermore, the precision values on the "Online" and "Grocery" datasets is 96.26 and 96.64, respectively.

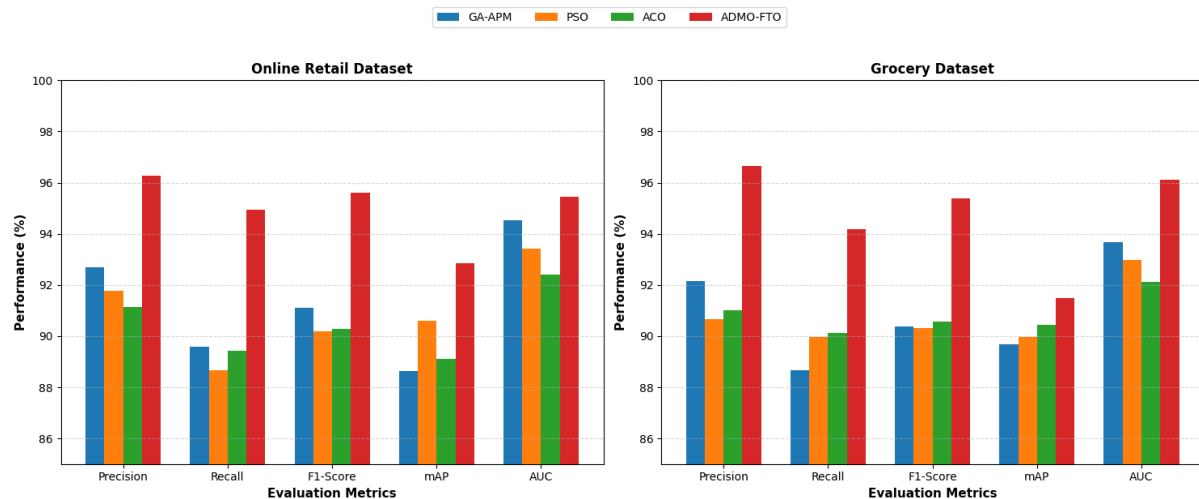


Figure 8: Graphical Illustration of ADMO-FTO vs Existing Techniques

The graphic illustration is evident in the ADMO-FTO framework (Figure 8) that demonstrates higher performance that is clearly visible in the figure. The proposed consistently maintained good performance than other existing performance on both the datasets.

COMPUTATIONAL COMPLEXITY

The section mainly discusses about the computational complexity based runtime and accuracy stability. In comparison to current metaheuristic algorithm, the computational complexity study assesses the effectiveness and scalability of the proposed ADMO-FTO

framework. This section shows how ADMO-FTO achieves near-linear time complexity and constrained space needs by analysing runtime increase and accuracy stability under different dataset sizes and sparsity levels. The investigation

verifies that the adaptive optimization technique and fuzzy weighting enhance accuracy while guaranteeing scalability for large-scale transaction datasets.

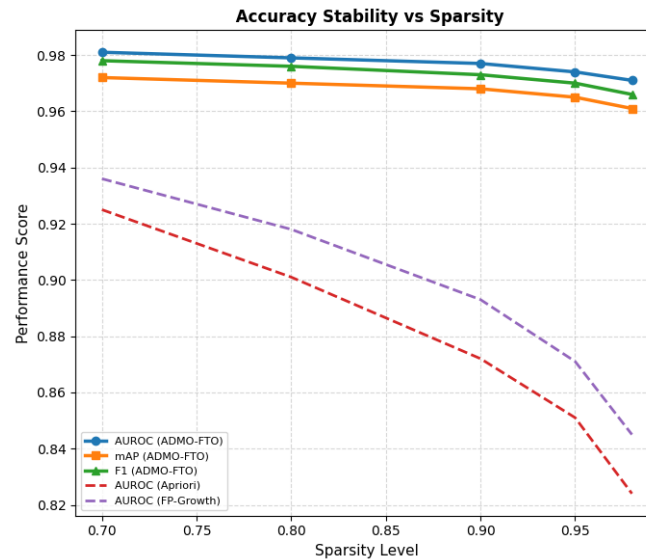


Figure 9: Graphical representation of sparsity based comparison

Figure 9 highlights the visual representation of ADMO-FTO performance under increasing sparsity level to ensure stability. In this case, when sparsity increases, the proposed ADMO-FTO approach outperforms GA-APM, PSO, and ACO in terms of F1-Score values while maintaining consistent performance. In sparse data conditions, ADMO-FTO consistently achieves higher AUC values that represents higher classification ability. In addition, the mAP results reveal that ADMO-FTO maintains higher accuracy while the performance of the existing approaches gradually decreases. In short, the findings show that ADMO-FTO performs better than the existing techniques and is more resilient to sparsity by retaining greater accuracy and prediction quality at various sparsity levels.

RUNTIME ANALYSIS

This section evaluates the execution efficiency of the proposed ADMO-FTO framework in comparison with GA-APM, PSO, and ACO. The results in Table 5, show that ADMO-FTO achieves near linear runtime growth as dataset size increases, while GA-APM and ACO exhibit significantly higher computational overhead. PSO converges faster on smaller datasets but becomes less stable with larger inputs. Overall, the runtime analysis confirms that ADMO-FTO delivers efficient scalability and reduced execution time, ensuring practical applicability for large-scale transaction datasets.

Table 5: Illustration of Runtime Analysis

Dataset	Transactions	Apriori (s)	FP-Growth(s)	ADMO-FTO (s)
OR-DS	5K	0.91	0.49	0.22
	10K	1.76	1.17	0.34
	15K	3.32	1.93	0.43
	20K	4.12	2.51	0.52
	25K	6.19	3.98	0.69
GR-DS	2k	0.22	0.19	0.09
	4k	0.53	0.54	0.15
	6k	0.87	0.61	0.28
	8k	1.54	0.76	0.54
	9.8k	1.76	0.96	0.97
	2k	0.22	0.19	0.09

Table 5 demonstrates the performance of various PM techniques including Apriori and FP-Growth are evaluated with the proposed ADMO-FTO framework. These evaluation is based on different combination of transactions to understand the scalability while increasing the volume of transaction. This comparison shows that the ADMO-FTO performs well on different scales than the other two techniques. The obtained value of ADMO-FTO is lower than the existing techniques that shows faster performance than other existing techniques.

CONCLUSION

This work establishes ADMO-FTO as a robust and scalable framework for CBA. The proposed work integrates three primary contributions. Firstly, an adaptive multi-objective Dwarf Mongoose algorithm to jointly optimize accuracy, sparsity reduction, and scalability. Secondly, it incorporates a sparsity-density analysis to ensure accuracy stability in the presence of heterogeneous itemsets. Thirdly, it introduces a flexible weighting mechanism based on Item Occurrence Analysis (IOA), capable of adaptively balancing frequent and rare items. This integrated technique effectively addresses the limitations of traditional algorithms, ensuring both computational efficiency and accuracy stability under sparsity. Comparative results show that ADMO-FTO consistently achieves superior performance across precision (96.26 (online) and 96.64 (grocery)), recall (94.95 (online) and 94.19 (grocery)), F1-Score (95.60% (online) and 95.40% (grocery)), AUC (95.45(online) and 96.12(grocery)), and mAP, while maintaining lower runtime growth than Apriori and FP-Growth. These findings confirm that ADMO-FTO is accurate and efficient, achieving a fair trade-off between predictive performance and computational complexity. To improve privacy, distributed scalability, and adaptation across various datasets, future developments will investigate cross-domain applications and integration with federated learning.

ACKNOWLEDGEMENT

We acknowledge the support of the PG & Research Department of Computer Science, Jamal Mohamed College (Autonomous), Trichy for providing computational resources. This research was not supported by any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Source of support: None

Conflict of interest: None

REFERENCES

- Agarwal, V., Gupta, R. K., & Tiwari, A. (2025). Multi-phase recommender framework for pattern warehousing using evolutionary optimization. *IEEE Access*, 13, 181925–181943. <https://doi.org/10.1109/ACCESS.2025.3622481>
- Chen, A. H.-L., & Gunawan, S. (2023). Enhancing retail transactions: A data-driven recommendation using modified RFM analysis and association rules mining. *Applied Sciences*, 13(18), 10057. <https://doi.org/10.3390/app131810057>
- Chen, D. (2015). Online Retail [Dataset]. UCI Machine Learning Repository. <https://doi.org/10.24432/C5BW33>.
- Chien, P.-Y., & Cadeaux, J. (2026). How assortment size, distinctive allocations, and local market density affect convenience store transaction performance. *Journal of Retailing and Consumer Services*, 88, 104492. <https://doi.org/10.1016/j.jretconser.2025.104492>
- Choi, J., & Park, H. Y. (2025). Usage complementarity vs. basket co-occurrence: Discount depth reliance in digitally personalized product recommendations. *Journal of Retailing*, 101(2), 177–196. <https://doi.org/10.1016/j.jretai.2025.01.006>
- D. Chen, S. L. Sain, and K. Guo, “Data mining for the online retail industry: A case study of RFM model-based customer segmentation using data mining,” *Journal of Database Marketing & Customer Strategy Management*, vol. 19, no. 3, pp. 197–208, 2012.
- Dogan, O. (2023). A recommendation system in e-commerce with profit-support fuzzy association rule mining (P-FARM). *Journal of Theoretical and Applied Electronic Commerce Research*, 18(2), 831–847. <https://doi.org/10.3390/jtaer18020043>
- Dogan, O., Kem, F.C. & Oztaysi, B. Fuzzy association rule mining approach to identify e-commerce product association considering sales amount. *Complex Intell. Syst.* **8**, 1551–1560 (2022). <https://doi.org/10.1007/s40747-021-00607-3>
- Fouad, M. A., Hussein, W., Rady, S., & Yu, P. S. (2021). An efficient approach for mining reliable high utility patterns. *IEEE Access*, 10, 1419–1431. <https://doi.org/10.1109/ACCESS.2021.3139028>
- Hahsler, M., Gruen, B., & Hornik, K. (2025). *Groceries dataset*. In *arules: Mining association rules and frequent itemsets* (R package). Comprehensive R Archive Network (CRAN). <https://CRAN.R-project.org/package=arules>

- Hahsler, M., Hornik, K., Reutterer, T. (2006). Implications of Probabilistic Data Modeling for Mining Association Rules. In: Spiliopoulou, M., Kruse, R., Borgelt, C., Nürnberger, A., Gaul, W. (eds) From Data and Information Analysis to Knowledge Engineering, Studies in Classification, Data Analysis, and Knowledge Organization. Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-31314-1_73
- Huang, P., Cheng, W., Chen, J., Chung, W., Chen, Y., & Lin, K.W. (2021). A Distributed Method for Fast Mining Frequent Patterns From Big Data. *IEEE Access*, 9, 135144-135159.
- Kallay, P., Dan Mihoc, T. Comparative Analysis of Frequent Pattern Mining Algorithms. *Acta Univ. Sapientiae Inform.* 17, 8 (2025). <https://doi.org/10.1007/s44427-025-00008-1>
- Karulkar, Y., Srivastava, S., Nandwana, R. et al. The Growing Complexity of Consumer Choices: Unravelling Consumer Patterns with K-Means and Fuzzy Logic. *J Stat Theory Appl* 24, 1165–1195 (2025). <https://doi.org/10.1007/s44199-025-00143-w>
- Kumar, S., & Mohbey, K. K. (2022). A review on big data based parallel and distributed approaches of pattern mining. *Journal of King Saud University – Computer and Information Sciences*, 34(5), 1639–1662. <https://doi.org/10.1016/j.jksuci.2019.09.006>
- Nakano, S. (2023). Customer demand concentration in online grocery retailing: Differences between online and physical store shopping baskets. *Electronic Commerce Research and Applications*, 62, 101336. <https://doi.org/10.1016/j.elerap.2023.101336>
- Smaili, M. Y., & Hachimi, H. (2023). New RFM-D classification model for improving customer analysis and response prediction. *Ain Shams Engineering Journal*, 14(12), 102254. <https://doi.org/10.1016/j.asej.2023.102254>
- Suchacka, G., Chodak, G. Using association rules to assess purchase probability in online stores. *Inf Syst E-Bus Manage* 15, 751–780 (2017). <https://doi.org/10.1007/s10257-016-0329-4>
- Sun, X., Ngueilbaye, A., Luo, K., Cai, Y., Wu, D., & Huang, J. Z. (2024). A scalable and flexible basket analysis system for big transaction data in Spark. *Information Processing & Management*, 61(2), 103577. <https://doi.org/10.1016/j.ipm.2023.103577>
- Surana, I., Belavagi, M.C. & Ramyashree KALEFormer a transformer based model integrating KMeans clustering and LDA for market basket analysis. *Discov Computing* 29, 224 (2026). <https://doi.org/10.1007/s10791-026-10103-2>
- Tomáš Pitka & Jozef Bucko & Stanislav Krajčí & Ondrej Křídlo & Ján Guniš & Lubomír Šnajder & Lubomír Antoni & Peter Eliaš, 2025. "Time analysis of online consumer behavior by decision trees, GUHA association rules, and formal concept analysis," *Journal of Marketing Analytics*, Palgrave Macmillan, vol. 13(1), pages 29–52, March.
- Trasierras, A. M., Luna, J. M., Fournier-Viger, P., & Ventura, S. (2024). Data heterogeneity's impact on the performance of frequent itemset mining algorithms. *Information Sciences*, 678, 120981. <https://doi.org/10.1016/j.ins.2024.120981>
- Wan, X., Han, X. Efficient Top-k Frequent Itemset Mining on Massive Data. *Data Sci. Eng.* 9, 177–203 (2024). <https://doi.org/10.1007/s41019-024-00241-2>
- Wu, S. An association rule-based approach for frequent item mining of multi-stage access data. *Discov Computing* 28, 139 (2025). <https://doi.org/10.1007/s10791-025-09644-9>
- Yang, W., Han, M., Dai, Z. et al. A high utility itemsets mining algorithm based on co-evolution. *Sci Rep* 16, 566 (2026). <https://doi.org/10.1038/s41598-025-30206-y>
- Zhang, H., & Ju, Y. (2026). Analysis of customer shopping behavior on E-commerce platform supported by fuzzy clustering algorithm. *Discover Computing*, 29(1), 281. <https://doi.org/10.1007/s10791-026-10151-8>
- Zhao, W., Tian, C., Wang, X., Wang, D., & Zhang, W. (2026). MEU-miner: Bitset-guided multi-dimensional high-utility sequential pattern mining with parallelism and adaptive top-k. *Knowledge-Based Systems*, 337, Article 15314. <https://doi.org/10.1016/j.knosys.2026.115314>
- Zhu, Z., Lin, Z., Chen, L., Dong, H., Gao, Y., Liang, X., & Deng, J. (2023). Correlation knowledge extraction based on data mining for distribution network planning. *Global Energy Interconnection*, 6(4), 485–492. <https://doi.org/10.1016/j.gloi.2023.08.008>