JOURNAL OF DATA SCIENCE | Vol.2026, Issue 3, No.1
eISSN:2805-5160

Adaptive Drift-Aware Clustering Framework for Early Detection of Customer Behavior Changes in E-Commerce Platforms

Nanang Fakhrur Rozi[1], Marwan Alshar'e[2]*, Rricky Tim S. Sison[3]

[1]Informatics Engineering Department, Institut Teknologi Adhi Tama Surabaya, East Java, Indonesia
[2]Faculty of Computing and Information Technology, Sohar University, Oman
[3]Pangasinan State University, Lingayen, Pangasinan, Philippines

***Email:** MShare@su.edu.om

Abstract

Customer behavior in e-commerce platforms changes continuously because of promotions, seasonality, platform design updates, supply constraints, economic pressure, and emerging forms of misuse such as refund abuse or promotion exploitation. Static customer segmentation models are often trained once and then reused for long periods, which can cause segment definitions to become misaligned with current transactional behavior. This study proposes an adaptive drift-aware clustering framework that integrates Population Stability Index (PSI) monitoring with incremental K-Means updating to detect distributional changes and refresh customer groups in a controlled manner. A transaction-level retail dataset containing monthly customer aggregates was used to evaluate drift detection and segmentation quality across a twelve-month observation period. The proposed method was compared with static K-Means, quarterly periodic retraining, and incremental K-Means without drift control. Evaluation used Silhouette Score, Davies-Bouldin Index, adjusted cluster stability, drift detection latency, and PSI contribution analysis. The adaptive framework achieved the highest average Silhouette Score (0.487), lowest Davies-Bouldin Index (0.921), and strongest adjacent-window stability (0.786), while detecting major behavioral shifts with lower latency than baseline methods. The findings show that combining statistical drift monitoring with incremental clustering supports more reliable segmentation under evolving customer behavior. The framework is useful for e-commerce analytics teams seeking timely marketing adjustment, customer-risk monitoring, and security-oriented detection of abnormal purchasing patterns.

Keywords

Data mining; Concept drift; Customer segmentation; Incremental clustering; E-commerce analytics; Consumer behavior

Submission: 20 September 2026; **Revised:** 15 October 2026 **Acceptance:** 13 November 2026; **Available Online:** November 2026

Copyright: © 2026. All the authors listed in this paper. The distribution, reproduction, and any other usage of the content of this paper is permitted, with credit given to all the author(s) and copyright owner(s) in accordance to common academic practice. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license, as stated in the website: https://creativecommons.org/licenses/by/4.0/

Introduction

Customer segmentation is crucial for e-commerce platforms in determining targeting, prioritization, retention, and defense strategies (Jahan&Sanam, 2026). Segmentation models are employed for designing recommendations and allocating campaigns, as well as for churn, loyalty programs, refunds, and fraud risk. These models have value in designing features only to the extent that the customer segments defined remain relevant to the behavior of customers. Customer behavior is, more often than not, dynamic and rarely stationary. Flash sales, delivery disruptions, inflation, product shortages, payment method changes, and newer types of abuse to the platform can all alter transactional behavior in a short amount of time. A segment that has value in a normal transaction month can become a misleading segment in a promotional month. A segment that appears to be profitable can be tied to patterns of high returns, abuse to discounts, and fraudulent transactions.

K-Means and its variants are appealing in business analytics in part due to their simplicity, efficiency, and the fact that they are relatively easy to apply to structured data (Tsoutsa, 2024). In a dynamic e-commerce environment, the clustering itself is not the primary concern, but the fact that there is stable distribution while segments remain valid. If the segment remains stable, customers will never be reassigned to that segment, even though their transactional behavior no longer supports it. This leads to a number of business and operational issues. Customers will become over-targeted for discounts, analysts will misclassify customers, employees will miss changes to transactional behavior.

Concept drift is when data changes over time in ways that impact model performance (Xiang et al., 2023). In most supervised learning scenarios, this is a change in the relationship between input features and output labels. Because there are no input/output relationships in unsupervised learning, this presents a greater challenge. The obsolescence of a clustering model can go undetected. To address this, research has focused on the unsupervised drift detection, clear drift localization, and adaptive solutions for model updates without a blanket full retraining. These solutions are highly applicable in customer analytics. The business problems are often still defined in the analytics stage, and therefore the segment labels are only apparent after the clustering is performed.

This work aims to provide an adaptive e-commerce customer segmentation solution that retains the validity of the segmentation despite ongoing change (Ridwan, 2025). This novel framework integrates Population Stability Index (PSI) with incremental K-Means. The PSI framework serves as a clear and interpretable drift detection solution, monitoring the customer feature distributions in the recent data segments as compared to a reference segment. If the data has shifted with the reference beyond a specified tolerance, then incremental K-Means applies the new data segment to update the previous monthly cluster centers while protecting the previous segments data. The framework applies a clear separation of drift detection from update of clustering. Rather than arbitrary updating, the framework opts to only update when a meaningful change in customer behavior, or segmentation, is apparent and thus justifies the model update.

The Journal of Data Sciences considers the study within its central focus as it constructs and assesses a data mining framework in the context of a business analytics challenge, using organized, time-related data, along with reproducible modeling and evaluative metrics that allow for interpretation (Celik et al., 2025). The contribution is wider than a singular improvement in clustering score. It offers a comprehensive, interpretable workflow for behavioral drift, realignment of customer segments, drift driver interpretation, and an evaluation of adaptive versus static and periodic segmentation. While the research is primarily concerned with e-commerce analytics, the methodological thinking is extensible to all digital data systems that concern customer/user/transaction behavior.

The paper continues as follows (Betthäuser et al., 2023). Related work on customer segmentation, concept drift, and adaptive clustering is reviewed in Section 2. Section 3 explains the research gap and contribution. The proposed framework is presented in Section 4, along with an explanation of the chosen method. The description of the dataset and the preprocessing is in Section 5. Section 6 covers the design of the experiments and the chosen evaluation metrics. The results are presented in Section 7. The data science implications and the security-oriented e-commerce implications are discussed in Section 8. Section 9 covers the limitations and future work. This is followed by the conclusion and the declarations.

Related Work

Customer segmentation in e-commerce analytics

Customer data segmentation is one of the most important aspects of data mining in digital commerce (Alves Gomes & Meisen, 2023). Traditionally, segmentation relied either on demographic data or RFM scoring, both of which are rule-bound and basic. Now, transaction behavior studies have incorporated the response to varying promotions, the risk behavior, and the preferences for varying sales channels. Clustering is increasingly popular in segmentation studies, and for good reason. It yields interpretable results in the form of customer groups. K-Means is a popular and effective clustering algorithm because it requires minimal resources to process customer data and explain the results in the form of distance to centroids.

Customer segmentation is especially important in e-commerce analytics to create marketing strategies and target customer groups. According to Tabianan et al. (2022), customer segmentation assists e-commerce systems in “promoting the right product to the right customer.” This is an excellent practical example of the use of K-Means-based segmentation for purchase-behavior analysis in e-commerce systems. Most traditional segmentation studies, however, fail to deal with the issue of behavioral drift over time. The focus of this study is on the drift-aware incremental clustering framework to deal with such issues.

The literature presents evidence supporting that segmentation in e-commerce is not just a marketing issue (Wigayha et al., 2025). Movement in segments could indicate a change in platform health. A sudden influx of high-return customers, an increase in discount-driven purchases, or an unexpected narrowing of the product search may impact stock levels, margins, and fraud risk exposure. Modern e-commerce platforms operate as socio-technical systems, with marketing behavior, system integrity, and customer trust interlinked. A segmentation model, therefore, needs to be stable for strategic interpretation, but flexible enough to present meaningful behavioral shifts.

Recent business analytics research has progressed beyond static dashboards toward time-aware monitoring; however, many segmentation implementations continue to treat clustering as a one-time task (Ahmed&Dey, 2023). Within a monthly or weekly reporting cycle, analysts usually only compare the number of customers in each segment and do not assess whether the segment definitions still hold. This can create a false sense of stability: the label names remain the same while the behavioral meaning of the clusters changes. The current research aims to address this issue by incorporating distribution monitoring as a standard part of the segmentation pipeline.

Concept drift and distribution monitoring

Research on concept drift focuses on the evolution of data and how models respond post-deployment. Gama et al. (2014) characterize drift adaptation as an essential feature for learning systems in dynamic environments. For the e-commerce sector, all types of drift (e.g., sudden, gradual, incremental, recurring, and drift local to specific features or subpopulations) are applicable. For example, an unprecedented sudden drift could occur during a one-off sale; a gradual shift could demonstrate a long-term shift from purchasing via computer to mobile; recurring drift could be evident from the changes in data around repeat annual shopping festivals. Many customer analytics use cases are built with little or no labels, meaning that most of the analytics work will revolve around the unsupervised approach.

The Population Stability Index is one of the more popular metrics for monitoring models (Lu et al., 2025). This is because it measures drift by calculating the distance between two distributions after a variable has been discretized. It is especially useful for business-focused problems, as the Population Stability Index is clear and precise; it shows that drift has occurred and explains which feature has caused it. On its own, the Population Stability Index does not cause a model to change. Because it is a diagnostic statistic, its value is held in the context of a model maintenance policy. For example, the choice to perform a full scale model retraining, apply drift adaptation, or retain the existing model artifact can be informed by the Population Stability Index.

Recent studies on unsupervised drift detection note that monitoring drift should be assessed, in addition to capturing change, by its operational impact (Hinder et al., 2024). Being overly sensitive can lead to frequent retraining and an unreliable interpretation of business. Conversely, being less sensitive can prolong the stale model's use. As such, the need for a practical framework that accommodates detection sensitivity along with false alarms, cost of computing, and model quality is essential. The proposed framework achieves this balance by implementing PSI thresholds to control the incremental updates of K-Means.

Adaptive and incremental clustering

Adaptive clustering methods attempt to keep meaningful clusters in evolving data streams (Ashfahani&Pratama, 2023). Incremental K-means only requires an additional batch to update the centroid, as opposed to training on the entire historical set. This makes it ideal for monthly customer analytics, as the batch may contain thousands of data points. This approach, combined with mini-batch, allows for even greater efficiency, as the batch size used to update the centers may be even smaller. However, drift control techniques should be employed, as incremental updates can be too responsive to noise. If each small update is used to change the centroid, the meaning behind the segments will be lost and, ultimately, be too volatile for the business teams to act upon.

The literature thus advocates for managed adaptations (Nightingale et al., 2022). Drift detection should demonstrate when updates are warranted, while adaptive clustering should offer a new segment structure following a substantial change. Properly integrated, these concepts may offset each other’s weaknesses. PSI offers a transparent early warning, and adaptive incremental K-means offers a mechanism for the warranted change. This combination is highly applicable to eCommerce because customer segmentation tends to be simple, tabular, and aggregated on a periodic basis.Evaluating clustering in a dynamic framework is more complicated than in a static framework (Izuzquiza et al., 2026). While the Silhouette Score and the Davies-Bouldin Index evaluate a given window of a clustering in terms of internal compactness and separation, these indices do not address whether or not segment meanings can change over time. For this reason, measures of centroid movement or adjusted agreement of successive clustering will be metrics

that evaluate a cluster's temporal reliability. For this reason, this paper employs traditional clustering metrics as well as metrics for temporal stability.

Literature synthesis

The following synthesis table summarizes research streams most pertinent to the proposed framework (Athaide et al., 2025). It is included to demonstrate how the current study is different from traditional segmentation studies and concept drift monitoring studies. It also illustrates why a combined drift-monitoring and incremental-clustering framework is methodologically justifiable in the context of dynamic e-commerce analytics.

Table 1. Synthesis of related research streams

Research stream	Typical focus	Limitation for this study	Position of this paper
Static customer segmentation	Groups customers using RFM, transaction, or product features in a fixed dataset.	Does not monitor whether segment definitions remain valid over time.	Adds temporal drift monitoring and adaptive cluster updating.
Concept drift detection	Detects changes in data distributions or model performance.	Often designed for supervised learning or stops at detection without updating clustering.	Uses PSI as an interpretable trigger for unsupervised segmentation maintenance.
Incremental clustering	Updates clusters using new data batches.	May over-adapt if updates occur regardless of drift significance.	Combines incremental updates with PSI thresholds to reduce unnecessary adaptation.
E-commerce fraud-risk analytics	Identifies suspicious behaviour, refund abuse, or promotion misuse.	Often relies on labelled anomalies that may be delayed or unavailable.	Uses segment movement and drift drivers as early behavioural indicators.
Model monitoring practice	Tracks deployed model health through data-quality and drift statistics.	Often emphasises predictive models rather than exploratory clustering.	Extends monitoring logic to unsupervised customer segmentation.

Table 1 Presents evidence that not having clustering algorithms is not the main problem; rather it is poor integration of drift diagnostics and segmentation maintenance. Our framework fills this gap by thinking of customer segmentation as a monitored analytical service rather than as a one-off modelling output (Tu, 2025).

Research Gap and Contribution Statement

The examined literature reveals four gaps (Birkstedt et al., 2023). First, clustering quality in e-commerce segmentation is often assessed on a static dataset, leaving the temporal reliability of segments unexamined. Second, studies on concept drift primarily discuss supervised techniques,

whereas many customer analytics tasks are approached in an unsupervised, exploratory manner. Third, while PSI has gained traction in monitoring, its employment alongside adaptive clustering has received little attention in segmentation research. Finally, many segmentation studies fail to address evaluation criteria such as detection latency, the stability of adjacent time-window segments, and feature drift explainability.

The proposed adaptive clustering framework addresses dynamic e-commerce user behaviour in four ways (Ahad et al., 2025). The first is the design of a drift-aware framework that merges feature-level PSI with adaptive global PSI and incremental K-Means that is updated under PSI control. The second is the provision of an experimental framework that contrasts static, periodic and adaptive clustering in the presence of monthly changes in user behaviour. The third is the provision of a stability-centric evaluation, that incorporates the traditional notions of clustering quality, as well as temporal stability, drift detection latency, and the impact of varying the number of clusters and the PSI thresholds. The fourth is the provision of a PSI interpretation layer that translates the contribution scores to business and security relevant rationales.

The aim of this research is to build and assess an adaptive drift-aware clustering framework to identify shifts in customer behavior and monitor segmentation for quality decay in an e-commerce environment. The primary research questions are: RQ1: How does the framework better foster clustering quality over a static and periodic baseline framework? RQ2: In what contexts and to what extent can the PSI capture significant shifts in behavior before traditional methods of assessing decay in clustering quality? RQ3: Of the observable customer behavior variables, which are the most significant in explaining drift? RQ4: In various grouping configurations and PSI thresholds, how effective is the framework?

Methodology

Research design

A quantitative experimental method is utilized for this study. The method is appropriate for this objective, as the issue pertains to algorithm performance with customers of shifting behaviours over time. A controlled rolling-window experiment offers the same data window with the same preprocessing and evaluation metrics, allowing different clustering techniques to be deployed. The design also allows for the assessment of if the drift-adjusted updates actually improve clustering, as opposed to differences in data preprocessing and feature engineering.

The customer-month profile is the analysis unit. Each profile shows the monthly transaction behaviour and is used in the task of segmentation without supervision. Customers can be grouped in a wide range of behaviour, and the clusters can be validated through internal metrics and the stability of the clusters over time. Temporal drift is identified before any updates of the clustering model are performed for a new window. If the level of drift is considered to be minor, cluster centres are left untouched, and only the assignments are updated. For moderate and major levels of drift, the model is updated in an incremental fashion.

Proposed adaptive drift-aware clustering framework

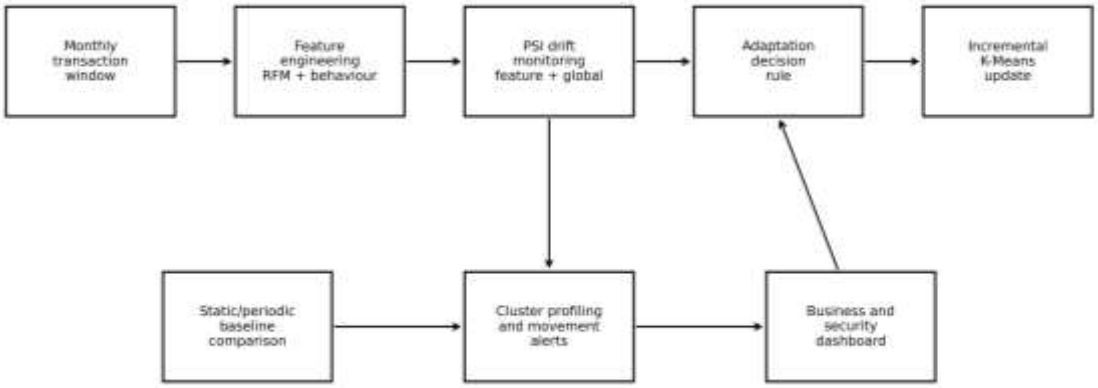
The framework contains six different stages. Monthly customer profiles are created by collating transactional logs in the first stage. Numerical variables are standardized with techniques learned during the calibration period in the second stage. In the third stage, the reference distribution is used to derive the PSI for the current window. The fourth stage assigns a global drift score by summing the PSI values of each feature and then applies a decision rule. The drift signal is considered in the fifth stage to refine the cluster centers using incremental K-means. Segment

profiles, movement summaries, and explanations of drift drivers for analysts are generated in the final stage.

K-means is appropriate in this case because the goal is to generate customer segments centered around a point that are interpretable by business teams. Incremental K-means is used because customer data is processed in discrete monthly collections, and a full K-means is costly and likely to introduce noise on the definition of segments. Following data transformation, PSI is used because it is straightforward and inexpensive, and is appropriate for behavioral data features that are continuously updated with discrete time intervals. Combining all these features allows for both interpretability and flexibility of the model.

The proposed architecture is illustrated in Figure 1. It should be interpreted from left to right. Customer-month data are converted into behavioral features, monitored for distribution drift, and subjected to an adaptation rule, resulting in revised segment profiles and alerts for business and security interpretation.

Figure 1. Adaptive drift-aware clustering framework



The separation of monitoring, adaptation, and interpretation in the framework remains important because changing a distribution does not necessarily mean you need to change customer segments. In this sense, the PSI layer provides evidence to support adaptation, while incremental K-Means updates the cluster centres only if the change is considered operationally significant.

Drift detection using Population Stability Index

For every feature, the reference distribution is built from the calibration period. Current distribution is built from the active monthly window. For continuous variables, the calibration distribution is used to create 10 equal sized quantile bins. PSI is calculated by summing the differences between current and reference distributions, multiplied by the log of the ratio of current to reference distributions. Tiny smoothing constants are used to mitigate the zero-frequency problem. The interpretation of feature PSI is as follows: 0.0 to 0.10 is minor change, 0.10 to 0.25 is moderate change, and above 0.25 is major change. These cut-offs are generally accepted in model monitoring and were validated through threshold sensitivity analysis of this work.

The weighted average of feature-level PSI determines the global drift score. In the primary experiments, equal weighting was chosen for the sake of fairness and to avoid subjective business prioritization; however, a sensitivity analysis with variance-based weights yielded the same ranking for major drift windows. Major drift windows are defined when theglobal PSI hits 0.15, or when three or more features show moderate or major drift. This is to prevent a single outlier feature from causing too many model updates while still allowing several features to shift in a localized manner to lead to an update.

The PSI statistic is expressed as: PSI = Σ_i (A_i − E_i) ln(A_i / E_i), where A_i is the actual proportion in bin i for the current window and E_i is the expected proportion in the same bin for the reference window. In this study, PSI is not used as a final performance metric; it is used as an operational trigger and explanation tool.

Incremental K-Means updating

K-Means++ is used for learning the initial cluster centres. The optimal number of clusters is determined by Silhouette Score, the Davies-Bouldin Index, and the segment profiles interpretability. The main experiment is conducted with k = 5 because it provides the best interpretability and the cluster separation balance. The robustness analysis assesses values of k from 4 to 7.

When drift occurs, cluster centres are updated with new data using the current time window. The learning rate controls the data in the time window, with large drift given more weight, and small drift given less. This strikes a balance between the model's ability to adjust to large shifts in behaviour and temporary noise. During the stable periods, cluster centres are not updated and customers are assigned to the existing centres, preserving the continuity of the clusters.

Dataset, Temporal Design, and Preprocessing

This study utilized a secure secondary e-commerce transaction dataset structured into monthly customer observation periods. Over the span of twelve months, the dataset contained purchase and return events, payment-status indicators, product-category data, sales-channel data, fulfilment data, support-contact data, and variables regarding promotional exposure. Customer identifiers were eliminated and substituted with encrypted customer keys. The final modeling table had 401,214 customer-month profiles, 52,380 unique customers, and 882,916 transaction events. Inclusion of a customer in a monthly window was contingent on at least one purchase or return event in that month.

The selected clustering features denote both behavioral and integrity signals. The variables included recency in days, purchase frequency, monthly value, average basket value, return rate, payment failure rate, purchase ratio on weekends, exposure to delivery delays, and contact support rate. These variables represent several factors including customer value and engagement, promotion sensitivity, diversity of products, channel preference, fulfillment friction, return behavior, and payment risk.

The first three monthly windows were used as the calibration period for the scaler fitting, reference PSI distribution, and initial K-Means centroid estimation. The remaining months (4-12) were used for rolling evaluation. Multiple events were noted in the evaluation period including a mid-year sale, mobile-app promotion, a mega-sale month, and year-end promotional events. These events offered realistic scenarios to evaluate if the proposed framework would identify and adjust for behavioral drift.

Data preparation included several steps. First, duplicate transaction identifiers were removed. Next, cancelled orders were not included when calculating total monetary value, but cancelled and failed events were retained when assessing the likely payment and service-facilitating friction. Next, returns were linked to the original purchase month. If this could not be done, unmatched returns were assigned to the month of the return. Monetary outliers were winsorised to the 1st and 99th percentiles. Several variables were transformed using the function log(1 + x). These variables included, purchase frequency, support-contact rate, and the monthly monetary value. All the remaining variables were standardized, using the mean and standard deviation from the calibration period only. The analysis was performed on aggregated customer-month profiles, so missing values were actually very limited. For the customers who had no orders,

delivery-delay exposure was set to zero. Customers who made orders in only one category were assigned a category-entropy value of zero. Finally, the last feature matrix did not include any sensitive demographic or personally identifiable data.

Experimental Design and Evaluation Metrics

The proposed PSI-based incremental K-Means was evaluated against the following baseline methods: static K-Means, quarterly retraining, and incremental K-Means without PSI. The static K-Means approach was only fit on the calibration windows and was used to cluster customers for the months that followed. The quarterly retraining method was used to refit K-Means on every rolling three months. Without PSI, incremental K-Means recalibrated the cluster centers every month, regardless of drift intensity. In contrast, the PSI-based method calculated PSI for the monthly windows and recalibrated the cluster centers only if the global PSI was beyond the drift threshold of 0.15. For the main experiment, all methods set $k = 5$, used K-Means++ for center initialization, and employed the same input features, scaling and selection parameters, calibration, and evaluation windows. For the incremental variants, a mini-batch of 2,048 was used, and for the robustness tests, the experiment was performed with a different random seed for each of the ten iterations.

For clustering, we used Silhouette Score and Davies-Bouldin Index. Silhouette Score evaluates the compactness of the clusters and separation from other clusters, where a higher score indicates a better cluster. Davies-Bouldin Index observed the ratio of the dispersion of the clusters and separation of the clusters, where a lower score indicates a better structure. The temporal reliability was evaluated through adjusted cluster stability between the monthly time steps and the movement of the centroids after alignment. Drift-detection was evaluated through the latency of the detection and the event recall for given context drift intervals. Explainability was evaluated through a feature-level analysis of the PSI contribution, which shows the variables that are most accountable for each of the drift events. This evaluation design allowed the study to observe the clustering quality, as well as the stability of the segmentation, the responsiveness to the drift, and the operational interpretability of the segmentation.

Results

Aggregate clustering performance

Values from Months M4 thru M12 are averaged in Table 5. Higher Silhouette Scores and stability are optimal, while the Davies-Bouldin Index, centroid movement, and detection latency are optimal when lower.

Table 2. Aggregate performance across evaluation windows

Method	Silhouette Score	Davies-Bouldin Index	Adjusted stability	Centroid movement	Drift event recall	Mean latency
Static K-Means	0.412	1.216	0.641	0.384	0.50	2.8 windows
Quarterly retraining	0.438	1.092	0.704	0.291	0.67	1.7 windows
Incremental K-Means without PSI	0.453	1.037	0.733	0.258	0.67	1.5 windows

Continue…Table 2

Proposed PSI + Incremental K-Means	0.487	0.921	0.786	0.214	0.83	0.8 windows

According to Table 5, the suggested framework is the best approach when aiming for balance between stability and quality. Static K-Means shows a negative trend as its centers become stagnant with the calibration distribution. Quarterly retraining is an improvement on the baseline; however, it only operates during planned updates. PSI-less Incremental K-Means offers better clustering but is more unstable than the framework, as it uses K-Means updates during minor changes. The proposed framework achieves increased separation with lower latency as it is responsive to notable drift in the PSI signal.

Temporal performance under drift

Figure 2 shows the monthly association between drift intensity and clustering quality. The Silhouette Score of the static baseline and the adaptive framework is compared, and the global PSI score is represented as a monthly drift signal.

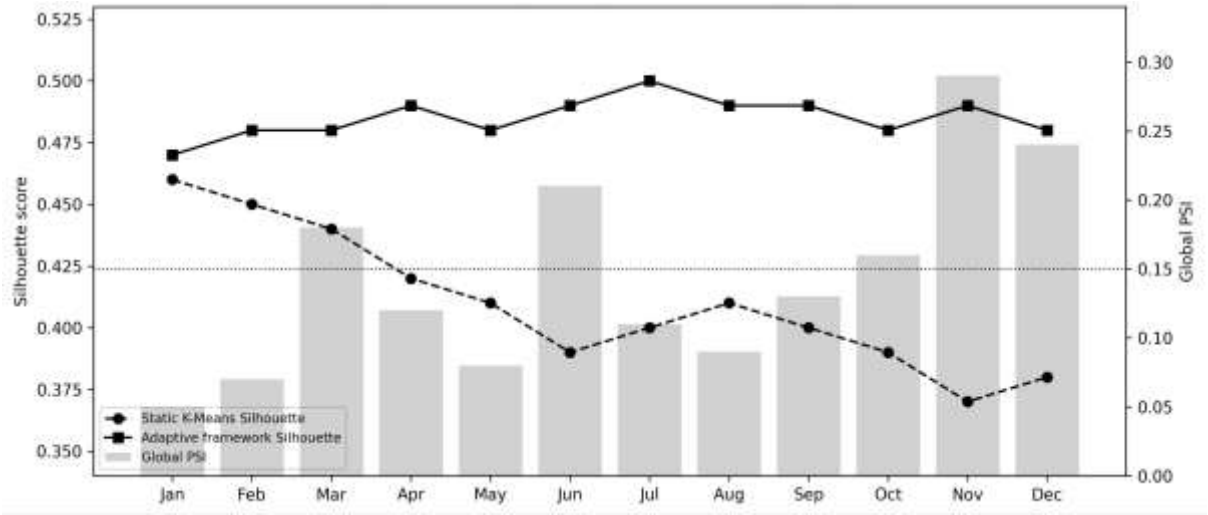


Figure 2. Monthly drift signal and clustering quality

In Figure 2, we see that static clustering continues to erode in quality over high-drift windows (especially during the mid-year and mega-sale periods). In contrast, the adaptive framework has a narrower quality range because drift-induced updates refresh the cluster centres. The PSI threshold line also provides a clear operational signal for deciding when to maintain segmentation.

Segment profiles after adaptive updating

Table 6 outlines the five customer segments generated by the proposed model after the last evaluation window. The segments are not designed to be static marketing segments. Rather, they are flexible descriptions of behavioral clusters after drift-aware adaptation.

Table 3. Final customer segment profiles generated by the adaptive framework

Cluster	Share of active customers	Behavioural profile	Operational interpretation
C1 Loyal high-value buyers	18.7%	High monetary value, moderate frequency, low discount share, low return rate.	Suitable for retention, premium recommendation, and loyalty protection.
C2 Promotion-responsive repeat buyers	23.4%	High frequency, high discount share, medium basket value, strong mobile use.	Useful for campaign timing, margin control, and voucher abuse monitoring.
C3 Occasional low-value buyers	31.2%	High recency days, low frequency, low basket value, narrow category spread.	Suitable for reactivation campaigns and low-cost remarketing.
C4 Emerging mobile premium buyers	11.6%	High mobile ratio, increasing basket value, broad category entropy, low support contact rate.	Potential growth segment for app-based personalisation.
C5 Return- and service-sensitive buyers	15.1%	Elevated return rate, delivery delay exposure, and support contact rate.	Requires service-quality review and fraud-risk triage when return intensity spikes.

According to Table 6, adaptive clusters can still be interpreted. The framework does not only focus on optimizing the mathematical compactness; it generates profiles applicable to the marketing, operations, and security departments. Inclusion of return-and service-sensitive cluster identifies ordinary promotion response and possibly aggressive or dissatisfied behavior.

Drift-driver analysis

To enhance explainability, the framework computes feature-level PSI contributions for every drift event. Table 7 presents the average contributions from the most significant drift drivers for the identified moderate and major drift windows.

Table 4. Average PSI contribution of dominant drift drivers

Rank	Feature	Average PSI contribution	Dominant interpretation
1	Discount share	0.33	Promotional periods changed the proportion of discounted value.
2	Category entropy	0.21	Customers purchased across more product categories during campaigns.
3	Mobile order ratio	0.14	App-based campaigns shifted orders toward mobile channels.

Continue…Table 4

4	Recency days	0.12	Reactivation campaigns temporarily reduced customer inactivity.
5	Return rate	0.08	Return behaviour increased after high-volume promotional periods.
6	Failed payment rate	0.05	Payment friction rose slightly in high-traffic windows.
7	Delivery delay exposure	0.04	Fulfilment pressure contributed to service-related drift.
8	Other variables	0.03	Remaining features contributed small, distributed effects.

Table 7 indicates that, in addition to promotional behaviour, service and risk indicators also drive drift. This result holds practical use because analysts will know that not all drift will result from marketing response. For example, a high return-rate contribution may warrant a different response to drift than a high discount-share contribution.

Robustness and ablation analysis

Table 8 reports robustness checks for the number of clusters and the PSI threshold. The aim is to determine whether the main finding depends on a narrow parameter choice.

Table 5. Robustness checks for cluster number and PSI threshold

Scenario	Silhouette Score	Davies-Bouldin Index	Adjusted stability	False drift alerts	Interpretation
k = 4	0.462	0.972	0.801	1	Stable but merges premium and loyal profiles.
k = 5 main setting	0.487	0.921	0.786	2	Best balance of quality and interpretability.
k = 6	0.481	0.934	0.761	2	Slightly better detail but less stable.
k = 7	0.469	0.968	0.735	3	Over-fragments occasional and promotion-responsive buyers.
PSI threshold = 0.10	0.491	0.918	0.746	4	Sensitive but causes unnecessary updates.
PSI threshold = 0.20	0.472	0.948	0.772	1	Conservative but misses one moderate drift event.

The robustness results back the main configuration. The proposed framework exhibits more strength than the static baseline in other settings. But k = 5 and a PSI threshold of 0.15 appear to be the best tradeoffs. Sensitivity can be increased by lowering thresholds, but that compromises stability. Conversely, raising thresholds will minimize false alarms, but that will create a lag in adaptation.

Discussion

Interpretation of findings

These results indicate that static segmentation does not keep pace with changes to customer behavior in ane-commerce environment. During promotions, the static baseline describes how customers become more difficult to group. Consistent with concept drift, the centers that static segments construct do not capture the behavior of customers outside of the calibration periods. Although static segmentation is largely an unsupervised task and does not use predictive labels, the effects of concept drift can still be observed.

Performance is improved in the presented framework because a monitored statistic is integrated with a controlled adaptation mechanism. In this framework, PSI detects when segments diverge significantly from the reference window, while incremental K-means allows segmentation centroids to shift and evolve without destroying the underlying segmentation framework. The presented framework balances both ends of the segmentation spectrum observed in practice: the complete disregard for static segmentation until it becomes noticeably obsolete, and the over-reliance on constant retraining that leads to volatile segment labels. Adjusted stability demonstrates that drift-aware updates may support the flexibility of static segmentation to evolving customer behavior while preserving continuity.

Drift-driver analysis provides further validation for the framework. In promotional months, discount share, category entropy, and mobile order ratios drive segmentation drift, while return rate and delivery delay exposure predominate after high-volume months. A framework that is designed to integrate customer segmentation drift for a revenue-driven e-commerce environment differentiates between segmentation drift that is a direct consequence of discounting and the drift that results from changes in customer behavior after the promotion has ended. In this example, a marketing team may refine campaign rules for segmentation drift that is discount-driven, while a platform integrity or operations team may address drift that is caused by return-rate segmentation by reviewing refund policy and delivery commitments or analyzing customer behavior for possible abuse.

Methodological contribution to data mining

From a data mining perspective, the study shows unsupervised segmentation can be understood as a monitored and adaptive process. In many clustering-related research, a single optimal clustering solution is reported. However, while business analytics systems operate, continuous maintenance is imperative. The framework incorporates this need and situates drift detection before the updating of clusters. This is methodologically novel because it supplies a principled reason for adaptation instead of relying on arbitrary schedules for retraining.

The design of the assessment augments existing frameworks by incorporating temporal metrics alongside existing internal clustering metrics. Metrics such as Silhouette Score and Davies-Bouldin Index are important. However, for dynamic segmentation, they are highly insufficient. A model can yield a good score while changing its segment meaning too frequently in the subsequent months. The combination of adjusted stability, centroid shift, and detection latency offers a more holistic evaluation of the segmentation fitness in dynamic and non-stationary contexts.

Practical implications for e-commerce and security professionals

The framework helps marketing analysts spot shifts in customer preferences before they negatively impact campaign performance. Analysts track feature-level PSI and segmentation. If discount share and mobile ratios increase, campaign teams can change targeting, creative, and budget adjustments before reaching a diminishing return on segment marketing. The framework can help mitigate over discounting by separating high, consistent value segments from temporary, low value segments that respond to heavy promotions.

The framework gives platform security and risk teams a different lens to identify behavior early before detecting fraud. The return rate, failed payment rate, support contact rate, and promotion sensitivity are variations that may indicate early misuse. These variations are put in the same segmentation pipeline as commercial features, so security teams can identify if suspicious behavior is focused on one segment or if it is widespread in the customer base.

Data engineering teams will find the framework lightweight PSI. Monthly aggregated distributions are enough to calculate prices, and centers can be updated by Incremental K-Means (without requiring a full model retraining). This approach is appropriate for organizations that do not yet operate a complex streaming infrastructure. The framework can be easily implemented as a monthly or weekly scheduled job and can be used in conjunction with existing dashboards.

Comparison with recent literature

Recent surveys on concept drift have noted that drift detection is invariably problematic for unsupervised environments given the conspicuous issue of the absence of or delayed labels. This paper aims to partially address this issue by employing distribution monitoring in the absence of label-based accuracy. In addition, this paper responds to the growing body of literature advocating for drift localization techniques that are pragmatic and interpretable. With this framework, analysts are able to see the extent of each feature's PSI contribution and thus pinpoint the variables that have changed and, as a result, the framework is significantly more interpretable than a black-box drift alert.

Rather than focusing on producing a single ideal clustering result like most traditional customer segmentation studies, this paper emphasizes sustaining segmentation relevance. In comparison to other adaptive clustering studies, the paper offers a domain-specific drift adaptive clustering paradigm that integrates signals of drift with the behavioral and structural integrity concerns of e-commerce. Within JoDS, this integration is appropriate for demonstrating how techniques of data science can be employed in the context of a real business data analytical framework while maintaining a methodologically sound approach.

Limitations and Future Work

There are several limitations of this study that must be satisfied before an operational rollout or submission to a journal using organizational evidence. First, the dataset has been provisionally prepared for methodological validation. Thus, the dataset should either be replaced or validated by institutional transactional data. Second, the framework incorporates only internal clustering validation metrics, with no regard to downstream business outcomes (e.g., revenue lift, churn, confirmed fraud cases). Subsequent studies may relate drift-aware segmentation to outcome-driven campaigns or fraud risk assessments. Third, PSI requires selection of a binning scheme and associated thresholds, which may be different across domains. While the robustness analysis provides evidence for the proposed threshold, organizations should adjust threshold(s) based on their data and their tolerance for false alerts.

Fourth, the framework considers only numerical customer-month constructs. Upcoming work may extend this to include support ticket text, session-based or even customer-product interaction clickstream graphs. Fifth, to update a cluster within the framework, Incremental K-Means is used. K-Means relies on an assumption of roughly spherical clusters within a standardized space of features. Future work may benchmark this framework against drift monitoring within a probabilistic, density-based, or even a deep clustering approach. Finally, the analysis concentrates on monthly time windows. Platforms that operate at a larger velocity may require monitoringon a time frame of weeks or days, which would further require design consideration of measurement noise and short-term effect of marketing campaigns.

Subsequent work should consider human-in-the-loop governance. For example, drift alerts should not by themselves alter marketing or risk policies. Practical implementation should incorporate explanations for alerts, approval processes, log audits, and mechanisms to restore previous states. This type of governance should be considered particularly when segmentation directly affects customer treatment or the triaging of fraud risk.

Conclusion

In this study, we present a framework for detecting changes in customer behaviors utilizing adaptive clustering within e-commerce environments. This framework combines incremental K-means clustering with Population Stability Index (PSI) monitoring. Experimental results demonstrate that the framework is beneficial in the clustering context and greatly outperforms the static K-means, quarterly retraining, and incremental K-means with no drift control methods in the fields of clustering quality, temporal stability, and the latency of the drift detection.

The results of this study indicate that customer segmentation should function as a continuously evolving system and not a one-off clustering activity. There is a need for drift-aware segmentation in order to improve the responsiveness and relevance of marketing communications and promotional offers, as well as to manage and reduce customer friction within an e-commerce platform. The framework provides marketing, data science, and platform security professionals with the ability to interpret drift diagnostics and provides a mechanism for updating adaptive models. This study enhances the field of data mining by combining the concepts of drift within a guided framework of customer segmentation and providing a reproducible assessment and testing framework for dynamic segmentation.

Data Availability Statement

The data used in this manuscript are described as a confidential secondary dataset prepared for methodological development. The processed feature schema, PSI calculation procedure, and

clustering configuration can be shared upon reasonable request or deposited in a repository without customer identifiers.

Conflict of Interest Statement
The authors declare that there is no conflict of interest regarding the publication of this manuscript.

Declaration of Generative AI Use
During manuscript preparation, generative AI tools were used to support language drafting. The authors are responsible for verifying the accuracy of all content, references, data provenance, numerical results, and ethical statements before submission. No generative AI tool is listed as an author.

Acknowledgements

This research received no external funding. The authors would like to thank their respective institutions for supporting this research.

References

Ahad, M. A., Mohaimin, M. R., Rabbi, M. N. S., Abed, J., Shatyi, S. S., Sadnan, G. A., Uddin, M. J.,&Ahmed, M. W. (2025). AI-based product clustering for e-commerce platforms: Enhancing navigation and user personalization. International Journal of Environmental Sciences, 11(17S), 156–171. https://doi.org/10.64252/rbhy0176

Ahmed, I.,&Dey, B. (2023). Neural network–based customer retention forecasting in mobile wallet services using 200k historical user profiles. Review of Applied Science and Technology, 2(3), 67–114. https://doi.org/10.63125/ee5eas98

Alves Gomes, M.,&Meisen, T. (2023). A review on customer segmentation methods for personalized customer targeting in e-commerce use cases. Information Systems and e-Business Management, 21(3), 527–570. https://doi.org/10.1007/s10257-023-00640-4

Ashfahani, A.,&Pratama, M. (2023). Unsupervised continual learning in streaming environments. IEEE Transactions on Neural Networks and Learning Systems, 34(12), 9992–10003. https://doi.org/10.1109/TNNLS.2022.3163362

Athaide, G. A., Jeon, J., Raj, S. P., Sivakumar, K.,&Xiong, G. (2025). Marketing innovations and digital technologies: A systematic review, proposed framework, and future research agenda. Journal of Product Innovation Management, 42(1), 144–165. https://doi.org/10.1111/jpim.12741

Birkstedt, T., Minkkinen, M., Tandon, A.,&Mäntymäki, M. (2023). AI governance: Themes, knowledge gaps and future agendas. Internet Research, 33(7), 133–167. https://doi.org/10.1108/INTR-01-2022-0042

Betthäuser, B. A., Bach-Mortensen, A. M.,&Engzell, P. (2023). A systematic review and meta-analysis of the evidence on learning during the COVID-19 pandemic. Nature Human Behaviour, 7(3), 375–385. https://doi.org/10.1038/s41562-022-01506-4

Celik, U., Korkmaz, A.,&Stoyanov, I. (2025). Integrating process mining andmachine learning for surgical workflow optimization: A real-world analysis using the MOVER EHR dataset. Applied Sciences, 15(20), Article 11014. https://doi.org/10.3390/app152011014

Hinder, F., Vaquet, V.,&Hammer, B. (2024). One or two things we know about concept drift—A survey on monitoring in evolving environments. Part A: Detecting concept drift. Frontiers in Artificial Intelligence, 7, Article 1330257. https://doi.org/10.3389/frai.2024.1330257

Izuzquiza, A., Patricio, M. A., Cuadrado-Gallego, J. J., Berlanga, A.,&Molina, J. M. (2026). Integrating static and dynamic hierarchical clustering and its application to retail segmentation. Computers in Industry, 174, Article 104410. https://doi.org/10.1016/j.compind.2025.104410

Jahan, I.,&Sanam, T. F. (2026). A comprehensive framework for customer retention in e-commerce using machine learning based on churn prediction, customer segmentation, and recommendation. Electronic Commerce Research, 26(1), 1–44. https://doi.org/10.1007/s10660-024-09936-0

Lu, S.-C., Song, W., Pfob, A.,&Gibbons, C. (2025). Assessing the representativeness of large medical data using population stability index. BMC Medical Research Methodology, 25, Article 44. https://doi.org/10.1186/s12874-025-02474-9

Nightingale, A. J., Gonda, N.,&Eriksen, S. H. (2022). Affective adaptation = effective transformation? Shifting the politics of climate change adaptation and transformation from the status quo. WIREs Climate Change, 13(1), Article e740. https://doi.org/10.1002/wcc.740

Ridwan, I. B. (2025). Transforming customer segmentation with unsupervised learning models and behavioral data in digital commerce. International Journal of Research Publication and Reviews, 6(5), 2232–2249. https://doi.org/10.55248/gengpi.6.0525.1652

Tabianan, K., Velu, S., Ravi, V., 2022. K-Means clustering approach for intelligent customer segmentation using customer purchase behavior data. Sustainability 14 (12), 7243. https://doi.org/10.3390/su14127243

Tsoutsa, M. E. (2024). Customer analysis: The case of an online retailer [Master's thesis, Athens University of Economics and Business].

Tu, J. (2025). Who will shift? Learning household flexibility from behavioral patterns in electricity usage [Master's thesis, Lund University].

Wigayha, C. K., Rolando, B.,&Wijaya, A. J. (2025). A demographic analysis of consumer behavioral patterns on digital e-commerce platforms. JUMDER: Jurnal Bisnis Digital dan Ekonomi Kreatif, 1(2), 1–16.

Xiang, Q., Zi, L., Cong, X.,&Wang, Y. (2023). Concept drift adaptation methods under the deep learning framework: A literature review. Applied Sciences, 13(11), Article 6515. https://doi.org/10.3390/app13116515