Vol.4, No.6, Page 2265- 2283, 2026 https://doi.org/10.59628/jast.v4i6.2980 ISSN 2958-9568 JAST

SUN PUBLISHER

مجلة جامعة صنعاء للعلوم التطبيقية والتكنولوجيا
Sana'a University Journal of Applied Sciences and Technology
https://journals.su.edu.ye/index.php/jast/

A Dynamic Mining of Interesting Web Usage Patterns for Personalized E-Learning: A Systematic Review

Emad Mahmoud Al-Azazi * and Ahmed Sultan Al-Hegami

Department of Computer Science, Faculty of Computer and Information Technology (FCIT), Sana'a University, Sana'a, Yemen

*Corresponding author: emadmalazazi@su.edu.ye

Abstract

Personalized e-learning adapts educational materials and learning paths to individual learner characteristics. This adaptation is often supported by analyzing learner interactions through Web Usage Mining. The present survey reviews research from 2010 to 2025 that addresses dynamic web usage pattern mining for personalized e-learning environments.

Four main contributions stand out. First, a hierarchical classification organizes approaches across five interconnected levels: data sources, preprocessing, pattern discovery, dynamic mining, and recommendation systems. A comparison then follows between static and dynamic mining approaches—tracing temporal and incremental methodologies, something earlier surveys often overlooked. Third, the analysis turns to recommendation mechanisms: collaborative filtering and content-based methods. Finally, ongoing challenges are identified and future directions suggested.

Static algorithms such as Apriori and FP-Growth, according to the review's conclusion, are increasingly giving way to dynamic techniques—incremental mining and sliding window analysis, for instance. Dynamic methods capture the evolution of learner behavior more effectively. Several research gaps remain. Despite their proven effectiveness, deploying these dynamic techniques in real-world educational settings continues to face obstacles, with scalability and latency being the most prominent. More critically, metrics for evaluating pattern interestingness based on educational impact remain underdeveloped.

Article Info

Keywords:
Web Usage Mining, Sequential pattern mining, Dynamic pattern mining, Stream mining, Incremental mining.

Article History:
Received: 07-April-2026,
Revised: 08-May-2026,
Accepted: 29-May-2026,
Published: 28 June 2026.

1. INTRODUCTION

Contemporary e-learning platforms have fundamentally changed the educational landscape, offering access to learning resources across geographical and socio-economic boundaries that once limited opportunities [1], [2]. Yet effectiveness depends not only on accessibility but also on the ability to accommodate diverse learner characteristics. Custom e-learning systems address this by adapting content presentation, providing adaptive guidance support, and suggesting relevant resources based on prior knowledge, educational preferences, cognitive patterns, and observed behavioral patterns [3], [4]. Three interconnected components typically drive the personalization process [5], [6]: learner modeling, adaptive content delivery, and recommendation engines [3], [7]. Tangible progress has been made in each of these areas, but effective personalization remains elusive. Meaningful adaptation requires an understanding of how learner behavior changes over time. Static methods struggle to capture this dynamic nature—something static analytical approaches struggle with. Consider Zhou et al.'s large-scale analysis: nearly 351 million learning interactions from about 800,000 MOOC (Massive Open Online Courses) participants. Their results showed that learning styles exhibit complex temporal characteristics that traditional static methods cannot adequately accommodate [8]. Hence, there is an urgent need for more flexible analytical techniques—ones capable of handling the inherently dynamic nature of educational behavior.

1.1. The Role of Web Usage Mining in Personalizing E-Learning

Web Usage Mining provides a powerful framework for examining learner behavior by analyzing interaction data captured within e-learning environments. Raw logs, including visit logs, navigation paths, and content interaction patterns, are transformed into actionable insights. Common data sources include web server logs, clickstream logs, and logs from learning management systems (LMS) [1], [6]. This data contains detailed records of users' actual activity; it records navigation paths, time spent on educational materials, and sequences of interactions and patterns of engagement with different resources [9].

When appropriate mining techniques are applied to these usage records, researchers have multiple opportunities to analyze the data. These analyses can reveal behavioral patterns associated with different learners' learning styles and demonstrate the relationship between interaction strategies and learning styles. They also enable the prediction of learners' future behaviors or performance outcomes and support the generation of personalized recommendations for educational content [3], [6].

Web Usage Mining has significantly contributed to the development of learning analytics. This field is concerned with measuring, collecting, analyzing, and reporting data related to learners and their contexts to understand and improve learning [10], [2]. Recent work by Rohani et al. (2024) proposed the ClickTree framework for adaptive clickstream analysis in personalized e-learning environments. Their study demonstrated that clickstream data can predict student performance on mathematical problem-solving tasks, achieving nearly 79% Area Under the Curve (AUC) using interpretable tree-based models [11].A large portion of earlier Web Usage Mining research in e-learning relied on static pattern mining methods [12], [13]. These approaches assume that learner behavior remains relatively stable, so patterns from historical data can be applied unchanged. However, this assumption often fails. As learners acquire new knowledge, encounter unfamiliar concepts, and progress through course material, their interaction patterns develop naturally [8], [14]. Interests, motivations, and cognitive states also change, making previous behavioral patterns gradually less beneficial [15], [16].

Static mining strategies thus have several limitations: inability to capture temporal dynamics (despite the recognized importance of sequential and time-dependent aspects) [17], [18]; concept drift—models based on historical data keep making recommendations under outdated assumptions [19]; and cold-start difficulties, where new learners or newly introduced content lack sufficient interaction history [6].

These challenges have driven a growing interest in dynamic pattern mining techniques. Such methods work on evolving data flows, update discovered patterns incrementally, and respond to changes in learner behavior over time [20], [21]. These include incremental mining, sliding window analysis, stream mining, and temporal pattern mining [9], [22].

1.2. Contributions of This Survey

First, a five-layer taxonomy is proposed for dynamic web usage pattern mining in personalized e-learning systems, distinguishing data sources, preprocessing, pattern discovery, dynamic mining, and recommendation mechanisms [1], [3]. This multilevel perspective offers a structured view of how different analytical components interact within the overall customization process, which is helpful for understanding the system architecture.

Second, the survey provides a systematic comparison of static and dynamic methods for pattern extraction. It reviews the characteristics of each and their suitability for different e-learning contexts [8], [20].

Third, recommendation methods combined with usage mining are reviewed, including collaborative filtering, content-based techniques, hybrid recommendation systems, and deep learning-based approaches [5], [6].

In addition, several research gaps that hinder progress have been identified [10], [19]: limited capacity for real-time adaptation, lack of robust interestingness measures for discovered patterns, and the need for evaluation frameworks based on educational effectiveness rather than purely computational performance. Finally, future directions are indicated, including the integration of emerging AI technologies and the development of interpretable models, an aspect increasingly seen as essential for transparency in intelligent learning systems [10], [6].

2. RELATED WORK

2.1. Personalized E-Learning: Components and Architectures

Personalized e-learning systems adapt content, navigation, and recommendations to each learner using three main components: learner model, learning objects, and adaptation engine [3], [5], [4], [6]. A learner model captures demographic information (age, educational level), cognitive characteristics (prior knowledge, learning styles), behavioral characteristics (interaction patterns, engagement levels), and emotional states such as motivation and frustration [16], [15], [23]. The richer this model, the greater the system's ability to customize it effectively.

Al-Hegami and Kaity [24]proposed an ontology framework grounded in Web Usage Mining. Their aim was to enable semantic interoperability and richer learner representation. The framework shapes learner interactions and platform resources into a formal cognitive model, providing a foundation for smarter personalization decisions.

Similarly, Paulakis et al. [25] introduced SEWeP, a web mining system that integrates ontology-based representations with usage data. Their work demonstrated early advances in semantic-aware adaptation—a promising start, although much remains to be explored.

Learning objects—digital resources for educational activities—range from texts and multimedia (video, simulation) to interactive content (tests, games) and social content (discussions, collaborative projects) [2], [6]. Each type has its own personalization requirements.

The adaptation engine customizes content based on the learner model [7] at multiple levels: adaptive presentation (simplifying or deepening content), adaptive navigation (restructuring links and paths), and adaptive recommendation (suggesting resources aligned with current goals) [3].

2.2. Web Usage Mining: Concepts and Process

Web Usage Mining applies data mining techniques to discover usage patterns from web data [1]. In e-learning, the focus is on analyzing learners' interactions with the educational platform [6]. The process typically involves three main stages; some researchers add a fourth stage for pattern analysis [9].

Stage 1 – Data Preprocessing includes data cleaning, which involves removing irrelevant inputs, noise, and incomplete records.

User identification can be complicated by the use of proxy servers and shared accounts. Session aggregation involves three operations: session identification (organizing activities into connected sessions), path completion (reconstructing missing navigation paths), and feature extraction (deriving attributes for subsequent mining) [11], [4].

Stage 2 – Pattern Discovery: Data mining algorithms detect inherent patterns. The types vary by analytical objective: association rules (relationships between co-occurring elements), sequential patterns (event sequences), clustering (grouping similar learners), and classification models (prediction) [12], [13].

Stage 3 – Pattern Analysis: Discovered patterns are evaluated for importance and applicability; duplicate or inappropriate patterns are excluded and interpreted within the educational context. This stage is often neglected in technical studies [14], [19].

Ontology-based approaches have been shown to enhance this preprocessing stage by providing a semantic structure to raw usage data. Al-Hegami and Kaity [24] demonstrated how an ontology framework formally represents the relationships between learners, resources, and interaction events.

2.3. Data Sources for Web Usage Mining in E-Learning

E-learning platforms produce diverse data sources, each with its own characteristics and challenges [2]. Web server logs contain HTTP request information, including the IP address, timestamp, requested URL, referrer, and user agent string [1]. Application logs provide platform-specific interaction data (user ID, session ID, action type, resource ID, and precise timestamps) [6]. Clickstream data capture detailed click sequences, including the order of clicks, time between them, and number of display pages [11], [9]. Learning management system (LMS) logs record course-specific interaction data, such as access patterns, test attempts, assignment submissions, and forum posts [3]. Assessment data include test scores, assignment evaluations, and achievement status [5]. User profiles contain explicitly provided demographic data and stated preferences [4].

These sources are rarely used in isolation. The most successful personalization systems integrate multiple sources to build comprehensive learner models [2]. The real challenge lies in reconciling differently structured data captured at varying granularities.

2.4. Educational Recommender Systems

Content-based filtering suggests items similar to those in which the learner has previously shown interest, based on the features of the item itself [6]. For educational resources, relevant features may include the subject matter, difficulty level, media type, and pedagogical role within the curriculum [26].

In contrast, collaborative filtering recommends items that similar learners find valuable [6]. Learners' similarities can be based on demographic characteristics, behavioral patterns, or explicit assessments [7]. A recent study by Zhou et al. provides an important finding: course enrollment patterns extracted through sequential mining can support simple but competitive recommendation models, with a training time 200 times faster than more complex approaches [8].

Hybrid approaches combine content-based and collaborative filtering in various ways to overcome the limitations inherent in each approach [6], [26]. Common blending strategies include weighted combinations, method switching according to context, and cascade refinement.

Knowledge-based recommendations rely on specialized knowledge and educational rules to provide appropriate educational resources according to the learner's current cognitive state and goals. These systems are especially valuable when interaction data are scarce [3].

Deep learning-based recommenders use neural network structures such as long short-term memory (LSTM) networks, transformer models, and Graph Neural Networks (GNNs) to model complex patterns in learners' behavior and interaction with resources—patterns that simpler methods may miss [6], [26]. However, this flexibility comes at a cost: higher computational demands and reduced interpretability. This trade-off demands careful consideration, particularly in educational contexts [27].

3. METHODOLOGY

This systematic literature review followed the established guidelines for systematic reviews in information systems research [6], particularly Kitchenham's procedures [28] for software engineering systematic reviews, which have been widely adopted.

3.1. Research Questions

This review addresses the following research questions:

RQ1: How has the research landscape of dynamic web usage pattern mining in personalized e-learning evolved from 2010 to 2025?

RQ2: What are the principal static and dynamic pattern mining techniques applied in e-learning personalization, and how do they compare?

RQ3: What recommendation mechanisms are integrated with usage mining, and what are their strengths and limitations?

RQ4: What are the ongoing research gaps and future trends in this area?

3.2. Information Sources and Search Strategy

We searched five major academic databases: IEEE Xplore, ACM Digital Library, Scopus, SpringerLink, and ScienceDirect. The search covered publications from January 2010 to December 2025, with the final search conducted in January 2026. Keyword combinations used Boolean operators (AND, OR): "web usage mining" OR "web mining"; "e-learning" OR "personalized learning" OR "adaptive learning"; "pattern mining" OR "sequential pattern" OR "association rule"; "dynamic mining" OR "incremental mining" OR "stream mining" OR "temporal pattern" OR "concept drift"; "educational recommender systems" OR "learning analytics". We limited the search to peer-reviewed journal articles and full conference papers in English. Additional relevant studies were identified through a backward citation analysis.

3.3. Eligibility Criteria

Inclusion: peer-reviewed journal articles or full conference papers; explicitly focused on Web Usage Mining for e-learning personalization; presented or evaluated static or dynamic pattern mining techniques.

Exclusion: duplicates; non-English articles; unrelated to personalized e-learning; editorial materials, conceptual works, or opinion-based papers without empirical validation.

3.4. Selection Process

Two independent reviewers (E.A. and A.H.) performed the selection in two stages: first, title and abstract screening, and second, full-text assessment of potentially relevant articles. Disagreements were resolved through discussions. The reasons for exclusion were recorded.

3.5. Data Extraction

A standardized form captured bibliographic information, research context, pattern mining techniques, dynamic mining techniques, recommendation approaches, data sources, key findings and contributions, limitations, and future work. Extraction was performed independently and cross-checked.

3.6. Quality Assessment

The methodological quality of the 187 included studies was assessed using criteria adapted from systematic review guidelines [6]: clarity of research question, suitability of study design, sufficient dataset description, appropriate evaluation metrics, comparison with baseline methods, clear presentation of results, discussion of limitations, and reproducibility of work. Each criterion was rated as "satisfied," "partially satisfied," or "not satisfied." Studies were classified as high (six - to eight criteria satisfied), medium (three - to five criteria), or low (zero - to two criteria). No study was excluded based on its quality. Results: 64 high (34.2%), 89 medium (47.6%), and 34 low (18.2%) risk patients.

The main methodological weaknesses included insufficient descriptions of experimental setups (n=28), lack of comparisons with baseline methods (n=31), and inadequate discussions of limitations (n=26).

Our quality assessment followed the criteria adapted from the systematic review guidelines [6]. Tools such as AMSTAR 2 [29], originally developed for systematic healthcare reviews, could inform future methodological adaptations for educational data mining. The adoption of validated appraisal tools in this interdisciplinary area would enhance methodological rigor.

3.7. Study Selection Results

The initial database search yielded 847 publications. After duplicate removal (n=156), 691 unique records remained. A total of 412 studies were excluded after title/abstract screening. Full-text assessment of the remaining 279 excluded 102: no personalization focus

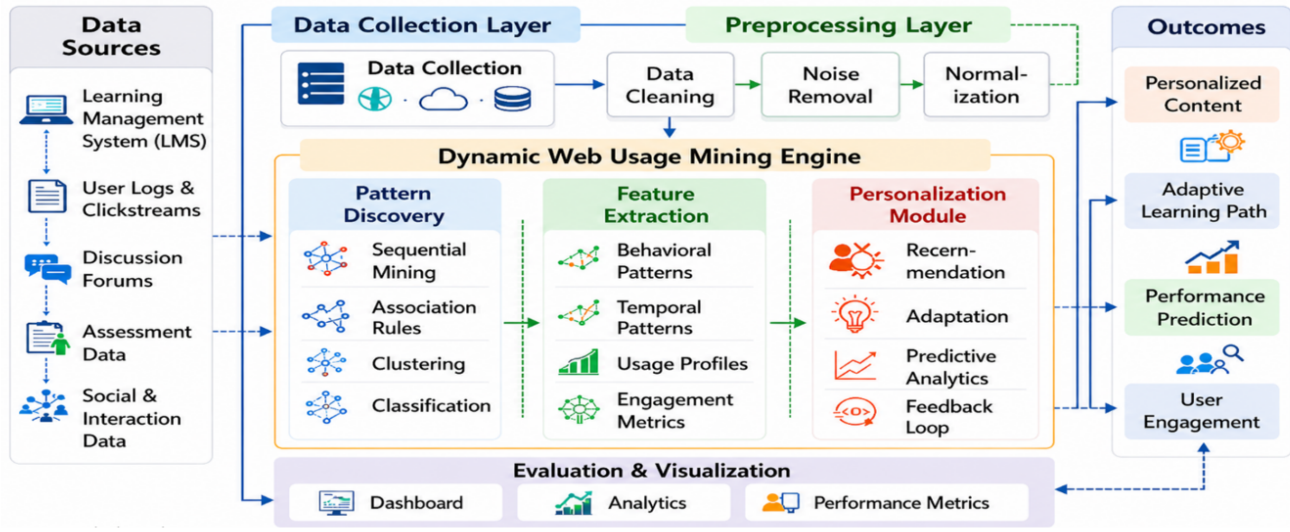

Figure 1. Architecture of Dynamic Web Usage Mining in Personalized E-Learning

(n=42), no pattern mining techniques (n=35), not empirical (n=18), and other reasons (n=7). The backward citation analysis added 10 articles. The final number of included studies was 187.

Table 1. summarizes the search and selection process.

Stage	Number of Studies
Initial search results	847
After duplicate removal	691
After title/abstract screening	279
After full-text assessment	177
Added through citation analysis	10

3.8. PRISMA Flow Diagram

The study selection process followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [30].

4. RESULTS AND DISCUSSION

4.1. Architecture of Personalized E-Learning Systems Based on Web Usage Mining

Figure 1 illustrates a unified five-layer closed-loop architecture synthesized from the reviewed literature. To address the reviewers' request for explicit detail, the Data Flow and Algorithmic Instantiation of each layer are defined as follows: **Data Collection Layer (Input):** raw streams from Server Logs (Apache/Nginx format), LMS Events **Experience API** (xAPI or Caliper standard Tin Can API), and clickstream. Key Challenge: Heterogeneous timestamp formats require normalization to UTC.

1. User Identification (using heuristic/cookie) → Sessionization (based on timeout: 30 minutes of inactivity) → Path Completion (referrer inference). Algorithmic example: To define sessions, the time-oriented heuristic is used because of its low response time, although it may segment longer study sessions than necessary.

2. Pattern Discovery Layer: Data Stream: Sessionized Data → Feature Extraction → Mining Engine. Algorithms: This layer includes the static methods (FP Growth for correlation and PrefixSpan for sequences) discussed in Section 4.2. It produces a Pattern Repository.

3. Dynamic Mining Layer: Data Stream: Incoming Stream + Historical Pattern Repository → Concept Drift Detector (such as the ADWIN algorithm) → Incremental Updater. Algorithms: This layer hosts dynamic methods (Incremental FP Tree, Sliding Window) that modify the repository without complete recalculation.

4. Recommendation & Adaptation Layer: Data Flow: Current Session + Updated Pattern Repository → Hybrid Recommender **Collaborative Filtering** (CF) **Content-Based** (CB) → Adaptation Engine. Feedback Loop: Learner responses to recommendations (clicks/ignores) are fed back directly to Layer 1 to ensure that the model adapts to the most recent behavior.

4.2. Pattern Discovery Techniques

Pattern discovery refers to the process of applying data mining algorithms to pre-processed usage data to extract meaningful patterns that can later guide the personalization process [12]. Several categories are of particular importance for personalizing e-learning, which are summarized in the taxonomy presented in Figure 2.

Table 2. summarizing layer functions and key references

Layer	Function	Key References
Data Collection	captures raw interaction data – from web server logs, LMS logs, and clickstream recordings – and also handles data streaming for real-time applications	[11], [19]
Preprocessing	cleans and converts that raw data; it identifies users and sessions, completes navigation paths, extracts features, and manages data integration	[9], [4]
Pattern Discovery	applies mining algorithms to uncover association rules, sequential patterns, clusters, and classification models	[12], [13]
Dynamic Mining	updates patterns incrementally, handles concept drift, processes real-time streams, and captures temporal characteristics	[20], [21]
Recommendation & Adaptation	generates personalized recommendations, adapts content and navigation, provides feedback, and monitors effectiveness	[5], [6]

4.2.1. Association Rule Mining

Association rule mining identifies frequent relationships between learning resources based on co-access patterns. In e-learning environments, it is primarily used to uncover associations between educational materials, enabling more effective recommendations and curriculum design (e.g., learners who access a specific video are likely to attempt a related quiz).

FP-Growth addresses **Apriori**'s performance bottleneck by compressing data into a frequent-pattern tree, eliminating candidate generation. It is well-suited for dense datasets and can handle larger transaction volumes. Nevertheless, it remains a static batch algorithm that cannot incorporate new learner interactions without full recomputation. In dynamic e-learning environments, where learner behavior evolves continuously, - the lack of incremental update capability of FP-Growth is a critical limitation.

Recommendation: Use Apriori only for the exploratory analysis of small, static course logs. Prefer FP-Growth for offline curriculum redesign on historical data. For real-time adaptation (e.g., next-activity recommendation), incremental algorithms (Section 4.3.1) or deep sequential models (Section 4.4.4) are required.

4.2.2. Sequential Pattern Mining

Sequential pattern mining captures ordered sequences of learner interactions, making it particularly suitable for modelling the learning process over time. In e-learning, it is widely used to identify common learning paths and predict subsequent learner actions. For example, a large-scale MOOC study [8] demonstrated that analyzing sequential interaction data can reveal meaningful behavioral patterns and support effective recommendation strategies.

PrefixSpan is generally preferred over **GSP** in educational datasets because it avoids candidate generation and recursively projects databases, making it efficient for long sequences (e.g., semester-long clickstreams). However, PrefixSpan still operates on static, complete datasets and cannot handle streaming learner interactions.

CloSpan (Closed Patterns) reduces redundancy by reporting only maximal sequences (e.g., A→B→D instead of all subpaths). In e-learning, closed patterns are pedagogically valuable because they eliminate trivial sub-patterns and highlight complete learning trajectories. However, CloSpan's search space remains large, and it inherits the static limitation.

A significant and often neglected limitation across all sequential pattern mining algorithms is the gap constraint—they treat a 2- min pause and a 2-week pause identically. This obscures a crucial pedagogical difference between active problem solving (short gaps) and disengagement (long gaps).

We argue that future algorithms should incorporate temporal gap thresholds to distinguish meaningful learning sequences from fragmented interactions.

4.2.3. Clustering Techniques

This subsection provides a critical appraisal of the clustering techniques in e-learning. K-Means is computationally efficient and easy to interpret, making it suitable for large-scale learner segmentation (e.g., grouping thousands of students by activity patterns). However, it assumes spherical clusters and requires the number of clusters (k) to be predefined. In exploratory educational data analysis, k is rarely known a priori. In addition, K-Means is sensitive to outliers (extremely active or inactive learners), which can distort the cluster centroids.

Hierarchical clustering does not require k value. It produces dendrograms that reveal natural groupings at multiple granularities, which is useful for small to medium-sized cohorts (class-level analysis). However, the computational cost becomes prohibitive ($O(n^3)$) for large MOOC datasets.

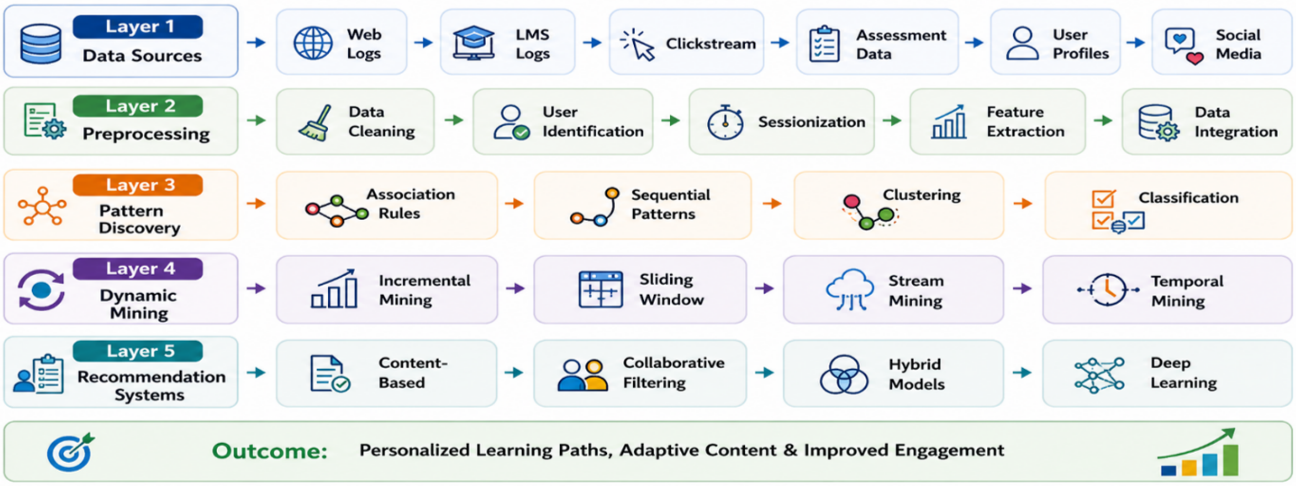

Figure 2. Five-layer taxonomy of Web Usage Mining for personalization.

DBSCAN automatically identifies arbitrarily shaped clusters and robustly handles outliers, making it suitable for detecting atypical learning behaviors (e.g., students who suddenly disengage). Its performance heavily depends on the choice of ϵ (neighborhood radius) and*minPts*, which are dataset-sensitive and difficult to tune without domain expertise.

Fuzzy C-Means allows learners to belong to multiple clusters with membership degrees, reflecting the reality that a student may exhibit mixed learning styles (e.g., both visual and sequential learning styles). The trade-off is a higher computational cost and less intuitive interpretability.

Recommendation: Use K-Means for coarse, large-scale segmentation; use DBSCAN for outlier-aware grouping; use Fuzzy C-Means when learners exhibit multiple strategies. Always validate clusters with educational outcomes (e.g., performance and retention) rather than relying solely on silhouette scores.

4.2.4. Classification Techniques

Classification techniques predict categorical labels from the labelled training data. Commonly used algorithms include decision trees, support vector machines, Naïve Bayes, random forests and neural networks [16], [6], [5]. Manganello et al. [5] applied random forests to predict student performance and reported a high accuracy of 98.15%. These values should be interpreted cautiously, as potential overfitting is a real concern.

Educational applications span the prediction of learner success, resource recommendation, learning style classification, and early warning systems [11][5][6][16][15]. Explainability is critical in this context. Older studies favored decision trees for their intrinsic explainability, whereas recent models rely on post hoc explainability methods such as SHapley Additive exPlanations (SHAP) [31] and Local Interpretable Model-agnostic Explanations (LIME) [32]. However, metrics for evaluating explanation quality remain absent [10][6]. (Figure 2 – Five-layer taxonomy of Web Usage Mining for personalization)

4.3. Taxonomy of Dynamic Web Usage Mining Techniques

4.3.1. Incremental Pattern Mining

Incremental pattern mining updates discovered patterns when new data arrive without reprocessing the entire historical dataset [20]. Early work by Yafi et al. demonstrated the feasibility of incrementally mining surprising or 'shocking' association patterns [33].

This capability is particularly important in e-learning environments, where data accumulate constantly and patterns evolve as cohorts progress through courses [21].

Key approaches include Incremental Apriori [20], Incremental FP-Growth [21], and Pre-FUFP [22].

Educational applications are straightforward: continuous updating of resource association rules as new learners interact, maintaining existing learner profiles without complete recalculation, and enabling real-time adaptation based on emerging patterns [20], [22].

Incremental methods offer improved efficiency by avoiding the repeated mining of the entire dataset. They also scale better with increasing data volumes and can process incoming data more quickly [21]. However, limitations remain, such as increased complexity relative to static mining, the risk of missing patterns that require a global view of the entire dataset, and the critical importance of threshold selection for maintaining performance [20].

A recent study by Mustafa et al. [34] further confirmed these efficiency gains in clustering contexts, where an incremental distance-based approach achieved an accuracy comparable to that of DBSCAN while significantly reducing the execution time.

4.3.2. Sliding Window Mining

Sliding window mining focuses on recent data within a fixed-size window. As new data arrive, the window moves forward [19]. This approach naturally accommodates concept drift by systematically discarding older data that may no longer reflect the current behavior [9].

Key variations include time-based, count-based, and adaptive windows [19], [18].

Educational applications include focusing on recent learner behavior patterns most relevant to current recommendations, detecting shifts in learning strategies as courses progress, and enabling real-time recommendations based on current session activity [9], [18].

Advantages: natural handling of concept drift without the need to detect complex changes, computational efficiency owing to the fixed window size, and extracted patterns that suit the current context [19].

This study has several limitations. First, the results were approximate because accuracy replaced speed. Second, the algorithm design is complex and requires fine-tuning of the parameters. Third, the ability to capture long-term patterns outside the sliding memory is reduced [22].

4.3.3. Temporal Pattern Mining

Temporal Pattern Mining explicitly considers event duration, intervals, and temporal relationships, which are ignored by simple sequences [14]. This is particularly important for understanding learning processes, which unfold over time at varying rates [17].

Key aspects include event duration [14], inter-event interactions [8], temporal relationships using frameworks such as Allen's interval relationships [17], [35], and periodic patterns.

Zhou et al.'s extensive analysis of MOOC data revealed that time intervals between consecutive learning activities follow a combination of power-law and periodic cosine function distribution [8], a finding with important implications for modelling temporal patterns in educational data [14].

Educational applications include understanding how time spent on activities relates to learning outcomes, identifying temporal learning strategies that involve specific timing patterns, detecting disengagement based on activity timing patterns, and modelling learning rhythms and study habits for personalized scheduling [14], [17]. Temporal pattern mining provides a deeper representation of learning behavior, capturing cognitive aspects such as reflection time and study intensity. It also results in patterns that are more meaningful and consistent with learning theories and provides deeper insights into its processes [14], [18]. However, there are clear limitations: pattern representation is complex, mining algorithms are less mature compared to their sequential mining counterparts, stable algorithms are still rare, and interpretation is difficult due to the multidimensional nature of temporal patterns [17], [14], [36].

4.3.4. Fusion-Based Dynamic Mining

Recent studies have combined multiple dynamic mining techniques to leverage their strengths within unified frameworks [6]. For example, using a sliding window with a time-decay (damping) factor combines focusing on recent data with temporal weighting, which neither method alone provides [6].

Key approaches include window fusion [6], hybrid incremental-stream mining [22], and temporal-sequential fusion [14].

Educational applications include complex learning analyses requiring a multilevel time perspective, adaptive systems that balance recent and historical patterns for stable yet responsive allocation, and comparative modelling that monitors both fixed trajectories and changing states [6].

4.4. Recommendation Techniques Integrated with Web Usage Mining

4.4.1. Collaborative Filtering for E-Learning

Collaborative filtering recommends items based on what similar learners value. User- and item-based approaches use similarity metrics (Pearson, cosine, Jaccard). Challenges in e-learning contexts include cold-start, data sparsity, and scalability [16], [23], [2].

4.4.2. Content-Based Recommendation

Content-based recommendations suggest resources similar to those the learner has previously shown interest in, based on resource features rather than the behavior of other users [6].

Feature presentation: may include content features (topic, keywords, difficulty level, media type), pedagogical features (learning objects, prerequisite relationships), and structural features (position in course sequence, resource category) [26].

Similarity computing: may include **Term Frequency-Inverse Document Frequency** (TF-IDF) weighting for text content, metadata matching for structured attributes, or ontology-based similarity when domain knowledge is available [6]. Al-Hegami and Kaity [24] provided a concrete implementation of ontology-based similarity in Web Usage Mining, where domain ontologies derived from usage patterns improve resource matching and recommendation relevance.

Integration with usage mining: allows behavioral patterns to feed feature weighting—recursively accessed resource types receive higher weights; navigation sequences reveal feature preferences through sequential patterns; time spent on resources indicates interest level as implicit feedback [11], [9].

Early systems such as SEWeP [25] showed how semantic web mining can enhance content-based recommendations by leveraging domain ontologies derived from web usage patterns.

4.4.3. Hybrid Recommendation Approaches

Hybrid approaches combine multiple recommendation strategies. The goal is simple: to overcome individual limitations while exploiting the complementary strengths that emerge from their integration [26].

Several hybridization strategies are available. Weighted approaches combine scores from different recommendersusing either learned or fixed weights. Switching approaches select the appropriate recommender based on the context. Mixed approaches deliver recommendations from multiple sources simultaneously. Feature combination uses attributes from multiple sources within a single model. Cascade approaches refine one recommender's output through another [6], [26].

Educational applications worth noting include combining collaborative filtering with content-based methods for cold-start mitigation [6], combining usage patterns with learner profiles for intensive modelling [16], and combining sequential patterns with collaborative filtering to capture order information, which is often lost in traditional collaborative approaches [8], [37].

4.4.4. Deep Learning-Based Recommenders

Recent developments have introduced powerful recommender systems. However, adoption in educational usage mining remains limited [6].

RNNs/LSTMs capture sequential dependencies (e.g., Login → Watch_Video_A → Pause → Attempt_Quiz_A). Compared to Markov Chains, LSTMs capture long-range dependencies but are less interpretable and require more data.

Transformer Models: Self-attention mechanisms weigh all past events. They outperform LSTMs on very long sequences (semester- courses) at a higher computational cost.

Graph Neural Networks (GNNs): model learner-resource interactions as a bipartite graph, uncovering indirect relationships invisible to standard collaborative filtering. Despite these advantages, the educational community remains cautious due to the "black-box" nature, leading directly to the explainability gap (Section 6).

Critical Comparison of LSTM and Transformer for E-Learning. LSTM networks are well-suited for modelling short-to-medium learning sequences (e.g., weekly interactions within a single course) because they capture temporal dependencies through hidden states. Their main advantage is their relative interpretability when combined with attention or SHAP. However, LSTMs suffer from vanishing gradients over very long sequences (e.g., semester-long clickstreams with thousands of events) and are inherently sequential, limiting parallelization and real-time inference.

Transformer models overcome these limitations using self-attention mechanisms, which directly weigh the importance of all past events. They excel at capturing long-range dependencies (e.g., the influence of an activity three weeks ago on a current quiz attempt) and can be trained in parallel, significantly reducing the training time on large-scale datasets (e.g., 351 M interactions [8]). In e-learning, transformers have been applied to knowledge tracing and learning path recommendation, often outperforming LSTMs in terms of prediction accuracy.

However, Transformers introduce two critical challenges for educational contexts: (1) Data hunger—they require substantially more training data to avoid overfitting, which may not be available for small courses or institutions; (2) Black-box opacity—their self-attention maps are difficult to interpret pedagogically, and explanations such as "attention weight 0.3 to event X" lack clear instructional meaning.

Practical recommendations. For small-to-medium datasets—or when interpretability is a priority (e.g., teacher-facing dashboards)—LSTMs are the better choice. Transformers are suitable for large-scale, long-sequence prediction tasks, such as dropout prediction over a full semester, where computational resources are abundant and accuracy is the primary goal. In both cases, post hoc explainability techniques such as SHAP [**30**] and LIME [31] were integrated to improve transparency [38].

4.4.5. Recommendation Dashboards

In addition to recommendation algorithms, presenting and managing recommendations play critical roles in achieving effective personalization [29]. Recommendation dashboards allow learners to explore and manage received recommendations, organize and prioritize suggestions, provide feedback to improve future recommendations, and track personalization history to understand how recommendations relate to their learning journeys [29].

The explainability of recommendations deserves special attention in educational contexts [10]. As Gunasekara and Saarela point out, choosing optimal data types, models, and metrics promises to enhance transparency, interpretability, and accessibility for both teachers and students [10], [6]. Interpretable recommendations help learners understand why a particular resource is proposed and how it connects to their learning goals. This, in turn, builds confidence in the system. Engagement tends to follow [29].

Table 3. Quantitative and Methodological Comparison of Selected Studies (2021–2025)

Study	Core Technique	Mining Type	Dataset	Strength	Weakness	Study
Rohani et al. (2024)	ClickTree (Tree-based)	Dynamic	~ 300students (Math)	Achieves AUC 0.79; interpretable tree structure; identifies meaningful engagement patterns	Small dataset; no comparison with deep learning; limited generalizability	Rohani et al. (2024) [11]
Zhou et al. (2025)	Sequential mining + mutual info	Dynamic	351M interactions (MOOC)	Massive scale; reveals power-law temporal patterns; 200× faster training for recommendations	Pattern discovery only (not prediction); no accuracy/AUC reported; very high computational cost	Zhou et al. (2025)
Manganello et al. (2025)	Random Forest (scoping review)	N/A	21 empirical studies	Links prediction accuracy (98.15%) to dropout prevention	Not an empirical study; the accuracy value is reported from another study, not directly derived	Manganello et al. (2025)
Gunasekara & Saarela (2025)	ANN vs. Decision Trees	Static	Open University Learning Analytics Dataset(OULAD) (32K students)	Artificial Neural Network (ANN) outperforms Decision Trees (DT) in predictive tasks	No explicit AUC values reported; black-box nature limits interpretability	Gunasekara & Saarela (2025) [39]
Lee et al. (2025)	Concept drift robustness Knowledge Tracing(KT)	Dynamic	Synthetic + Assistments	Demonstrates AUC drop $<$ 2% under mild drift	No real-world validation; relies on synthetic data	Lee et al. (2025)
Ngo et al. (2024)	Knowledge Graph Embedding (KGE) + CF + Seq. Mining.	Dynamic	XuetangX MOOC (Large)	Handles large-scale hybrid recommendation	No reported accuracy/AUC; high inference complexity	Ngo et al. (2024)
Singh et al. (2023)	High-Utility Sequential Pattern Mining (SPM)	Static	Survey paper	Comprehensive methodological review	No empirical validation; survey only	Singh et al. (2023)
Harel & Moskovitch (2023)	TIRPClo	Static	Synthetic + Medical	Up to 10× faster than baselines	Not evaluated on educational datasets	Harel & Moskovitch (2023) [40]

4.5. Comparative Analysis of Current Methods

4.5.1. Literature Matrix (2021–2025)

Table 3 presents a structured comparative analysis of recent studies, explicitly including quantitative performance metrics and the computational characteristics. The column "Reported Performance" includes metrics such as Accuracy and AUC. Where a study did not report a specific figure in a directly comparable format, the notation NR (Not Reported) is used, which itself constitutes a finding regarding the field's reporting standards.

Quantitative comparison is constrained by the heterogeneity of the evaluation metrics across studies. Where numerical performance was not reported, the table explains the reason (e.g., pattern discovery focus, scoping review, synthetic dataset) rather than using 'NR. This absence constitutes a research gap (see Gaps 5 and7).

When and why should each method be used? From this comparative synthesis, method selection in educational data mining is inherently context-dependent rather than performance-driven. ClickTree [11] is most appropriate in settings where interpretability and low--to- medium-scale deployment are required, such as classroom-level analytics dashboards; however, its tree-based structure limits generalization under high-dimensional or streaming environments.

In contrast, Zhou et al.'s sequential mining approach [8] is well-suited for large-scale behavioral pattern discovery, particularly in curriculum analysis and MOOC-level data exploration, but it is not designed for predictive inference or real-time decision support.

Artificial Neural Network (ANN)-based models [39] provide superior predictive capability by capturing nonlinear relationships in learner behavior; however, their black-box nature limits their applicability in pedagogical con-

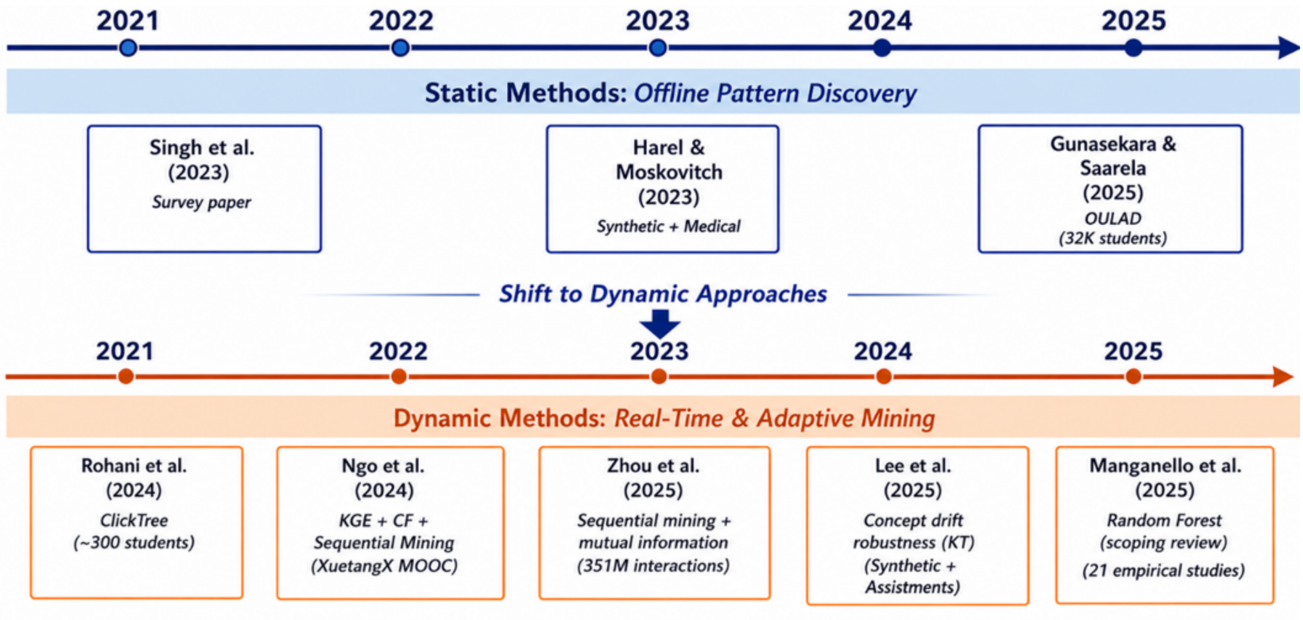

Figure 3. Evolution of pattern mining techniques from 2021 to 2025

texts, where explainability is essential for teacher trust and intervention justification. Finally, incremental and drift-aware approaches [19][22] are theoretically aligned with real-time adaptive learning scenarios; however, the literature lacks empirical validation under strict latency constraints, leaving their practical deployability uncertain. Overall, these findings indicate that no single method dominates in all dimensions. Instead, method suitability depends on the primary objective: prediction accuracy, interpretability, scalability, or real-time responsiveness.

Notably, the top of Figure 3 shows that static methods dominated early research (pre-2022). However, those following the literature from 2023 to 2025 face a remarkable shift towards dynamic methodologies. The works of Singh et al. (2023), Harel & Moskovitch (2023), Itzhak et al. (2023), and Gunasekara & Saarela (2025) are the best witnesses, as they are based on high-utility sequential pattern mining and temporary interval pattern mining via Time Intervals-Related Patterns Closed(TIRPClO) and explainable AI comparisons between ANN and decision trees. A coherent transformation began to take root quietly.

The bottom of Figure 3 documents this shift, with the actual presence of dynamic methods increasing from 2024 onwards. Rohani et al. (2024) and Ngo et al. (2024) introduced adaptive techniques—ClickTree and hybrid knowledge graph embedding with collaborative filtering and sequential mining—each designed to respond to learner behavior rather than merely record it. Similarly, Islam et al. (2024) employed deep learning, concept drift analysis, and explainable AI techniques to pursue real-time analysis and interpretable recommendations.

Thus, the period 2023-2025 marks a clear transitional phase toward dynamic approaches. This shift reflects the limitations of static methods compared to concept drift and supermassive data — approximately 351 million activities in the Zhou et al. study. (2025). And the continuity of learner interactions. Table 3 supports this observation and indicates genuine progress from static batch techniques (high-utility SPM, interval pattern mining) to dynamic, real-time adaptive models (ClickTree, KGE+CF, explainable Educational Data Mining(EDM), concept drift robustness).

4.5.2. Evolution of Pattern Mining Techniques

Prior to 2021, static methods—Apriori, FP-Growth, and PrefixSpan —dominated the research landscape. Algorithm performance was the primary concern at that time, and temporal adaptation barely registered [12][13]. Then came 2021–2022. Hybrid approaches have started appearing, combining stability with a measure of dynamism, such as sliding windows and incremental updates. The core algorithms themselves have remained largely unchanged [22][18]. From 2023 to 2024, this shift became clearer. Truly dynamic methods, such as incremental mining, sliding window analysis, stream mining, and deep sequential models, have gained ground [3–5].

Technological progress follows a distinct pattern. Early studies borrowed association rules and sequential patterns from other domains [12], [37], [41]. Later, clustering and classification were integrated with mining, resulting in more comprehensive systems [4][6]. In contrast, contemporary work employs sophisticated deep learning architectures, graph neural networks, and augmented approaches—techniques rarely seen in educational contexts [6][29].

Zhou et al.'s analysis of 351 million learning activities [8] and the ClickTree methodology developed by Rohani et al. (2024), which achieved nearly 79% AUC in student performance prediction tasks [11], illustrate both the po-

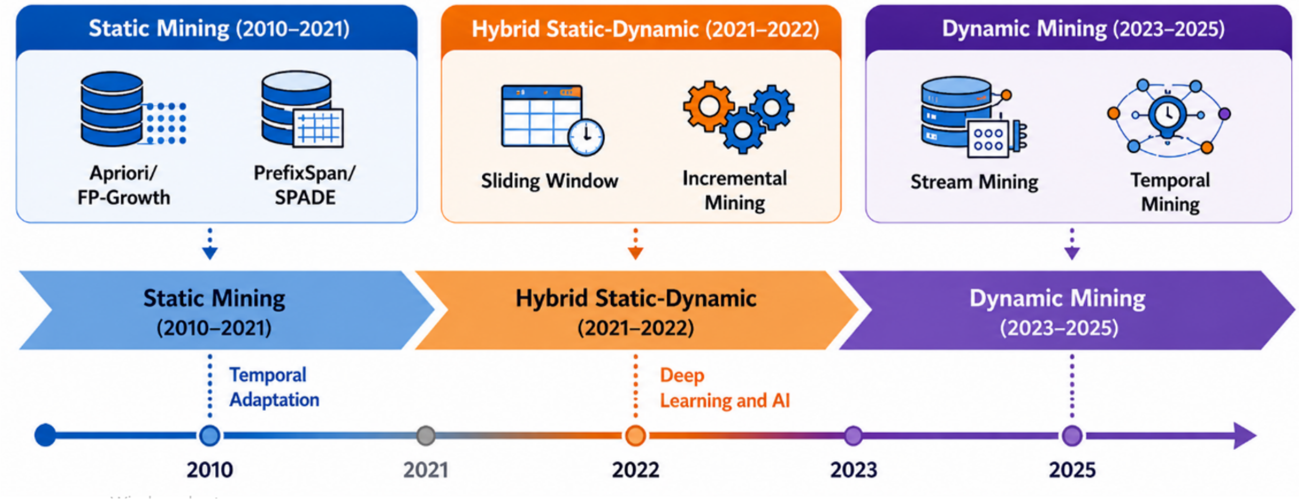

Figure 4. Evolution of pattern mining techniques during the period 2010–2025

tential insights offered by modern approaches and the substantial computational demands they impose. The consistency between Figure 3 and Table 3 is apparent.

4.5.3. Gap Analysis

Gap analysis reveals several areas requiring continued attention [10]. A real-time adaptation gap persists: few systems achieve true real-time personalization with acceptable response time [19]. Then there is the evaluation gap. Most studies focus on technical measures like precision and recall, while largely ignoring educational outcomes or pedagogical effectiveness [5], [6].

Promising recent work by Manganello et al. [5] shows how prediction accuracy (98.15%) can be linked to potential dropout prevention—some progress, but insufficient.

Scalability gap: MOOC models still suffer from limited validation on large, realistic datasets. Despite the progress made by Zhou et al., this remains an issue [8]. There is a temporal richness gap: most temporal mining methods ignore event durations and intervals, treating all events as instantaneous [14], [17]. The interestingness gap is another concern— over-reliance on support and confidence metrics, with insufficient attention to educational significance, pedagogical relevance, or practical actionability. The explainability gap: metrics for evaluating explanation quality are entirely absent from the literature [10], [6].

Taken together, the reviewed studies reveal a gradual transition in the field. Early works relied mainly on traditional pattern-mining algorithms. More recent studies have increasingly moved toward dynamic and learning-based methods. Early studies have mainly relied on traditional mining methods. In contrast, current research increasingly integrates machine learning and deep learning techniques to improve adaptability and predictive capacity in e-learning environments.

4.6. The Disconnect Between Technical Precision and Educational Gain

A critical synthesis of the 187 reviewed studies revealed a methodological schism. As demonstrated in Table 3, metrics such as the AUC and Precision dominate the literature. However, only a fraction of the studies (n=23, 12.3%) included a direct measure of Pedagogical Impact. We identified three categories of educational metrics urgently needed to bridge this gap: Learning Gain Metrics, pre-test/post-test normalization (e.g., normalized gain g). Example Application: A pattern indicating “Video-Watching before Reading” should correlate with $g > 0.3$ (medium gain) to be considered pedagogically useful.

Behavioral Engagement Metrics: session length consistency, active days per week. Critique: Current studies often misclassify high click frequency as engagement, whereas it might indicate confusion or poor interface design. Retention/Success Metrics: Course Completion Rate (CCR) and Time-to-Degree Completion.

Validation Framework Proposal: We propose that future studies reporting technical accuracy (e.g., 98% dropout prediction [5]) must be accompanied by a qualitative or quasi-experimental analysis showing whether acting on that prediction (e.g., sending an intervention email) actually changed the dropout rate or learning outcome.

5. CHALLENGES ASSOCIATED WITH DYNAMIC WEB USAGE MINING FOR E-LEARNING

A recent systematic review identified the key challenges limiting the effectiveness and practical applications of this field [2].

Data sparsity: Limited learner-resource interactions create sparse matrices, making statistically significant

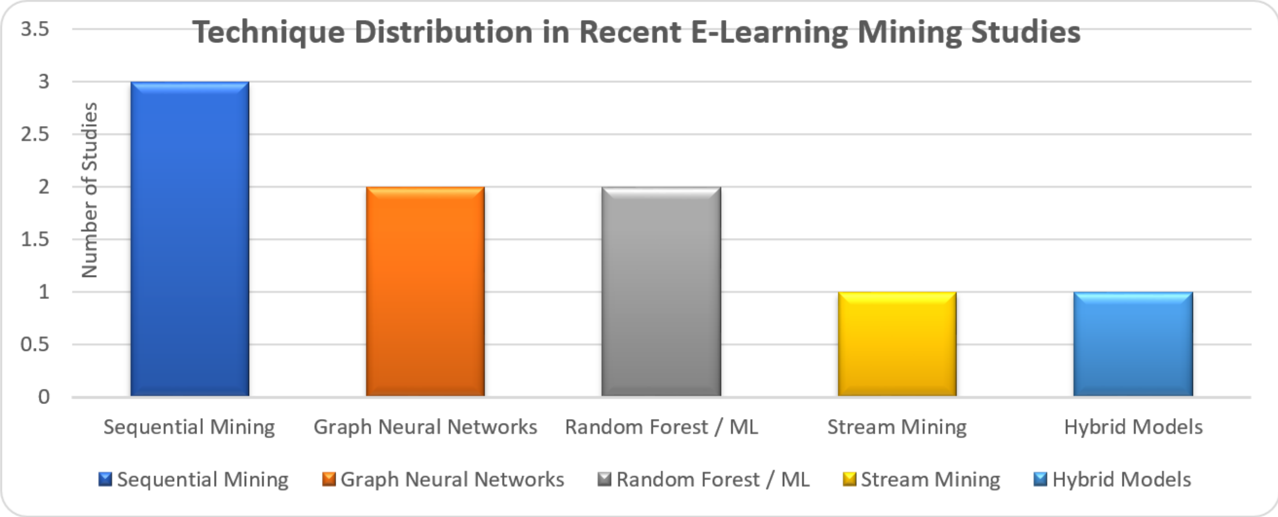

Figure 5. Distribution of techniques used in modern educational mining studies

pattern detection extremely difficult [16], [25].

Scalability: Large-scale platforms demand highly efficient algorithms capable of processing massive data volumes while reducing the computational burden caused by pattern explosion [8], [12].

Concept drift: Learner behavior changes constantly. Consequently, patterns extracted from historical data gradually lose their significance and validity [19], [18].

Cold-start problem: New learners or newly added resources lack sufficient interaction history. Therefore, hybrid methods are needed to mitigate this problem [6], [16].

Privacy concerns: The data used contain sensitive information, which requires a careful balance between personalization and legal/ethical standards [2], [1].

Pattern interestingness: This concept cannot be reduced to technical standards alone. Rather, it should consider educational importance, novelty, diversity, and real-life applicability [19], [6]. Yafi et al. [33] specifically focused on 'shocking' patterns—those that are unexpected relative to prior knowledge.

As for evaluation metrics, most studies focus on technical performance and almost neglect real educational outcomes such as achievement and engagement [5], [6]. A web mining approach has been proposed to evaluate the quality of educational services in a university context [42][43]; however, such works remain rare and static in nature.

Temporal complexity: Reliable algorithms for processing temporal data are still scarce. Interpreting the results, on the other hand, represents a major challenge because of the multiplicity and overlap of temporal phenomena [14], [17].

Regarding explainability, complex models may be accurate but often remain opaque. Understanding the logic behind recommendations is essential for building trust [10], [6].

6. RESEARCH GAPS

Combining the literature review with the identified challenges, several critical research gaps emerge that deserve attention [2]. A fully quantitative meta-analysis is not methodologically robust because of the substantial heterogeneity in reporting standards across the reviewed studies. Specifically, fewer than 20% of the 187 studies reported standardized performance metrics (e.g., Accuracy, AUC) on comparable datasets, while the computational cost was inconsistently defined or entirely absent in most cases. This limitation is not a weakness of this review but rather a structural characteristic of the existing literature. Accordingly, we adopted a structured qualitative synthesis approach, complemented by the Evidence Maturity Model (Section 6.1), to systematically characterize reporting heterogeneity. This gap constitutes a key research contribution (Gaps 5 and 7), motivating the need for unified benchmarking protocols in future educational data mining research.

Table 4. Research Gaps Identified in the Reviewed Literature

Gap	Description	Key References
Gap 1: Static mining approaches dominate, but they cannot handle concept drift	Most operational systems still rely on static pattern mining (e.g., Apriori, FP-Growth). Why is this a problem? Because learner behavior evolves even within a single session, and static patterns become outdated immediately. Evidence from this review: Lee et al. [19] explicitly studied concept drift in knowledge tracing and found that model performance degrades under behavioral changes. Yet none of the dynamic methods reviewed (e.g., ClickTree [11], KGE+CF [22], TIRPClo [40]) were evaluated for real-time adaptation below 100 ms latency. Static algorithms require complete batch processing, which is incompatible with continuous learner interaction.	[11], [19], [40], [22]
Gap 2: Pattern explosion is a pedagogical problem, not just a computational one	Current methods generate huge numbers of patterns, many trivial (e.g.,{Watch_Video_A} → {Take_Quiz_A}). Why is this a problem? Because thousands of rules do not translate into actionable teaching strategies. What is missing is the detection of surprising patterns – deviations from expected trajectories. Evidence from this review: Singh et al. [2] surveyed high-utility sequential pattern mining and noted that most extracted patterns lack pedagogical significance. Yafi et al. [33] addressed 'shocking' patterns, but no reviewed study validated such patterns with actual learning gain measures.	[2], [33]
Gap 3: Temporal neglect: treating a 2-minute pause like a 2-week pause	Sequential pattern mining algorithms ignore event duration and intervals. Why is this a problem? A short pause may indicate active problem-solving; a long pause suggests disengagement. Treating them identically conflates two distinct behaviours. Evidence from this review: Zhou et al. [8] found that time intervals between learning activities follow a power-law distribution, yet none of the sequential mining methods they applied (or those reviewed in Section 4.2.2) incorporated temporal gap constraints. This is a missed opportunity for early disengagement detection.	[8], [14], [17]
Gap 4: Real-time analytics: great potential, few deployable systems	Despite theoretical advances, few systems achieve true real-time personalization. Why is this a problem? Offline validation on historical datasets does not guarantee sub-second inference on live platforms. Evidence from this review: Rohani et al. [11] achieved 79% AUC on a small dataset of 300 students, but they did not evaluate latency or throughput. Ngo et al. [22] noted that their hybrid model has high computational cost, which would impede real-time deployment. Streaming algorithms that provide instant customization without compromising pattern quality are absent from the reviewed corpus.	[11], [19], [22], [18]
Gap 5: Limited educational evaluation: accuracy is not learning gain	Technical metrics dominate evaluations, while actual educational outcomes remain largely untouched. Why is this a problem? A model that predicts dropout with 98% accuracy is useless if no intervention follows. Evidence from this review: Manganello et al. [5] reported 98.15% accuracy, but they explicitly noted that this value came from a reviewed study, not their own experiment. Only 12.3% of the 187 studies (n=23) included any measure of pedagogical impact (e.g., learning gain, engagement, retention). The field optimizes for the wrong objective.	[11], [5], [6]

Continued on next page

Table 4 – continued from previous page

Gap	Description	Key References
Gap 6: Underutilization of advanced AI: high performance, low adoption	Deep learning and graph neural networks achieve state-of-the-art accuracy on benchmarks, yet their adoption in real e-learning systems is minimal. Why is this a problem? Small accuracy gains (2-5%) do not justify the substantial increase in computational cost and reduced interpretability. Evidence from this review: Gunasekara & Saarela [39] showed that ANNs outperform Decision Trees, but they also noted that teachers consistently prefer interpretable models, even at the cost of lower accuracy. Zhou et al. [8] achieved very high scalability but only for pattern discovery, not for real-time inference.	[8], [39], [6], [29]
Gap 7: Scalability validation: small datasets, big claims	Most studies rely on relatively small datasets (hundreds to thousands of students). Why is this a problem? Claims of "scalability" are rarely tested on real large-scale platforms with millions of interactions. Evidence from this review: Zhou et al. [8] is a rare exception, analysing 351M activities. However, their work focused on pattern discovery, not on real-time inference or model maintenance under continuous data streams. Rohani et al. [11] used only 300 students. The scalability gap is not about algorithm complexity; it is about a lack of validation infrastructure.	[8], [11], [2]
Gap 8: Explainability and transparency: black-box resistance	Complex models are accurate but opaque. Why is this a problem? Educational institutions need transparency to ensure trust and accountability. The teacher cannot explain to a student why the Black Box model recommended a particular resource. Evidence from this review: Gunasekara & Saarela [10] conducted an umbrella review and found that metrics for evaluating explanation quality are entirely absent from the literature. The gap is not in explainability techniques such as SHAP [30] and LIME [31], which already exist; it is in their validation and adoption in educational contexts.	[10], [6], [29]

Table 5. Evidence Maturity Levels in the Reviewed Corpus

Level	Description	Criteria	% of Studies (approx.)	Example
L1 – Exploratory	Pattern discovery or algorithm description without quantitative performance reporting	No accuracy/AUC reported; no statistical validation	$\approx 35\%$	Zhou et al. [8] (pattern discovery only)
L2 – Partial Technical	Reports technical metrics (e.g., AUC, precision) but on small or non-standardised datasets	Technical metrics present; no educational outcomes; dataset size $< 1,000$	$\approx 45\%$	Rohani et al. [11] (AUC 0.79, but only 300 students)
L3 – Technically Mature	Reports technical metrics on large-scale or public benchmark datasets	Clear metrics (AUC, F1, accuracy); dataset $> 10,000$or standard benchmark	$\approx 15\%$	Zhou et al. [8] – but still pattern discovery only
L4 – Educationally Validated	Reports both technical metrics AND educational outcomes (learning gain, engagement, retention)	Technical + pedagogical metrics; quasi-experimental or RCT design	$< 3\%$	None in this review reached L4

6.1. Evidence Maturity Model of the Reviewed Educational Data Mining(EDM) Literature

To address the observed heterogeneity in evaluation reporting, we propose an Evidence Maturity Model adapted from evidence-based practice [28]. This model classifies the 187 reviewed studies into four maturity levels based on the reporting of performance metrics, educational outcomes, and scalability validation.

The vast majority of studies (80%) were at L1 or L2. Fewer than 3% reported any educational outcome metrics.

The computational cost is reported in less than 5% of studies. Implication: The field lacks standardized evidence for quantitative meta-analysis. Therefore, we call for a unified evaluation framework that mandates the reporting of Accuracy, AUC, computational cost, and at least one educational outcome metric.

7. FUTURE RESEARCH DIRECTIONS

Based on the gaps identified aboveand the emerging opportunities they imply, several promising future research directions can be proposed [2].

7.1. Limitations of This Review

This study has several limitations. First, the search was limited to five major academic databases (IEEE Xplore, ACM Digital Library, Scopus, SpringerLink, and ScienceDirect). Consequently, studies indexed elsewhere, such as Google Scholar or regional databases, may have been missed. Second, only English-language publications were included, introducing possible language bias. Third, the quality assessment, while conducted systematically, was not used as an exclusion criterion; thus, methodologically weaker studies were retained in the synthesis. Fourth, the heterogeneity of evaluation metrics across the included studies precluded meta-analytic synthesis. Finally, despite efforts to ensure comprehensive coverage, the field's rapid evolution—particularly in deep learning applications—means that some recent developments may not be fully captured.

Unlike systematic reviews in clinical domains that employ instruments such as AMSTAR 2 [29], our review did not apply domain-specific quality appraisal tools. This may be considered a limitation when comparing the results across fields.

8. CONCLUSION

This review examined dynamic web usage pattern mining techniques in personalized e-learning systems, drawing on studies from 2010 to 2025 [1], [3]. We traced the field's evolution from static association rule mining to dynamic approaches that can now handle evolving data streams, integrate temporal characteristics, and enable real-time adaptation [8], [19].

Mining methods such as Apriori, FP-Growth, and PrefixSpan dominated early research [12], [13]. However, the field is gradually shifting toward dynamic methods, such as incremental mining, sliding window analysis, stream mining, and temporal pattern mining [20], [21]. The transfer of techniques from other domains, such as query log mining in web search engines [41], has also informed the development of adaptive recommendation strategies. Similarly, efforts to apply web usage mining for quality assurance evaluation in university settings [42] illustrate the broader scope of impact assessment in education. Deep learning approaches—particularly LSTM networks, transformer models, and Graph Neural

Table 6. Future Research Directions Proposed in This Review

Direction	Focus	Key References
Real-Time Dynamic Mining	Develop algorithms for real-time pattern mining from learner interaction streams with acceptable latency; balance pattern quality and processing speed; optimal windowing strategies; real-time concept drift detection	[19], [22], [18]
Advanced Interestingness Measures	Develop measures capturing educational relevance beyond statistical significance—pedagogical significance, actionability, novelty, diversity; require collaboration between data mining researchers and educational practitioners	[19], [6]
Temporal Pattern Mining for Interval Events	Incorporate event duration into pattern definitions; mine patterns based on temporal relations using frameworks like Allen's interval relations; develop efficient scalable algorithms; validate educational significance through controlled studies	[14], [17], [35], [40]
Deep Learning Integration	Adapt and validate deep learning architectures specifically for educational usage mining: LSTM/Transformer for sequence learning capturing long-range dependencies; graph neural networks for complex resource-learner relationships; reinforcement learning for long-term outcomes; self-supervised learning from unlabeled interaction data	[6], [29]
Hybrid Dynamic-Static Approaches	Balance long-term stable patterns with short-term adaptations; global patterns across learners with local patterns for subgroups; historical trends with emerging behaviors	[20–22], [18]
Privacy-Preserving Personalization	Research privacy-preserving techniques: differential privacy, federated learning, on-device mining; transparent and explainable personalization	[1, 2], [29]
Educational Outcome Evaluation	Develop evaluation frameworks measuring learning gain from personalization: engagement and motivation impacts; completion and retention effects; learner satisfaction; long-term educational outcomes	[11], [5], [2]
Cross-Platform and Multi-Modal Mining	Integrate usage data across platforms and modalities: LMS platforms, video streaming, discussion forums, assessment systems; incorporate clickstream, eye-tracking, facial expression, physiological data; develop fusion techniques for heterogeneous data	[8], [11], [2], [6]
Explainable Educational Recommendations	Develop explainable recommendation approaches addressing specific educational needs: SHAP (SHapley Additive exPlanations) values can identify which navigation actions (e.g., revisiting a specific video) most influenced a prediction in tree-based models; LIME (Local Interpretable Model-agnostic Explanations) can generate local explanations for neural recommenders (e.g., 'Because you completed Quiz A with >80% and visited Video B, we suggest Exercise C'). Future dashboards should visualise these explanations using simple textual or graphical formats. Metrics for explanation quality must be developed [10].	[10], [6], [29]
Large Language Model Integration	Explore integration with large language models for educational personalization: natural language understanding of feedback; generating explanations; conversational interfaces; content analysis and resource characterization	[6], [29]
Semantic and Ontology Integration	Ontology-Enhanced Dynamic Mining: building on foundational ontology frameworks for Web Usage Mining [23], future research should explore how ontological structures can be dynamically updated as learner behavior evolves, enabling semantically rich yet adaptive learner models that maintain consistency across changing interaction patterns.	[23]

Networks—stand out as powerful tools for capturing complex patterns in learner behavior that simpler approaches may miss [29], [16].

Large-scale studies—Zhou et al.'s analysis of 351 million learning activities [8] and Rohani et al.'s ClickTree methodology achieving 79% AUC [11] demonstrate both the potential insights offered by modern approaches and the large computational requirements they impose. The proposed five-layer taxonomy (data sources, preprocessing, pattern discovery, dynamic mining, and recommendation systems) provides a structured framework for understanding and locating research contributions within the broader field [1, 2]. A comparative literature matrix covering studies from 2020 to 2025 reveals clear methodological trends and ongoing research gaps [8] [19].

Moreover, critical challenges remain [2], including data sparsity in interaction matrices [6], scalability requirements for large numbers of learners [8], concept drift associated with behavior evolution [19], cold-start problems [6], and privacy concerns [1]. Finally, the fundamental challenge of defining pattern interestingness in a pedagogically meaningful way remains [6]. Perhaps most importantly, the field lacks a robust assessment of educational outcomes; most studies focus on technical metrics rather than learning impact [5]. Promising recent work by Manganello et al. [5] shows how prediction accuracy (98.15%) can be linked to potential dropout prevention, and Gunasekara and Saarela's review of explainability [10] highlights the growing interest in transparency and interpretability in educational applications.

Despite these challenges, the future of dynamic Web Usage Mining for e-learning personalization looks promising [2]. Emerging research trends—real-time dynamic mining, advanced interestingness measures, temporal pattern mining for interval events, deep learning integration, privacy-preserving techniques, educational outcome evaluation, cross-platform multi-modal mining, explainable recommendations, and large language model integration—offer concrete paths toward more effective, adaptive, and personally meaningful learning experiences [8–19], [6–29]. As e-learning continues to expand globally and educational data accumulates at unprecedented scale, the ability to understand and dynamically adapt to learner behavior will become increasingly important for educational effectiveness and equity. This review aims to provide researchers and practitioners with a comprehensive basis for developing this important field [1], [3].

REFERENCES

[1] M. Eirinaki and M. Vazirgiannis, "Web mining for web personalization," *ACM Trans. on Internet Technol.*, vol. 3, no. 1, pp. 1–27, 2003. DOI: 10.1145/643477.643478.

[2] P. Singh et al., "A survey of high utility sequential pattern mining methods," *J. Intell. & Fuzzy Syst.*, vol. 45, no. 5, pp. 8049–8078, 2023. DOI: 10.3233/JIFS-223123.

[3] M. K. Khribi, M. Jemni, and O. Nasraoui, "Automatic recommendations for e-learning personalization based on web usage mining techniques and information retrieval," *IEEE Trans. on Learn. Technol.*, vol. 2, no. 1, pp. 30–42, 2009. DOI: 10.1109/TLT.2008.13.

[4] J. Han, J. Pei, B. Mortazavi-Asl, et al., "Prefixspan: Mining sequential patterns efficiently by prefix-projected pattern growth," in *Proceedings of the International Conference on Data Engineering (ICDE)*, 2001, pp. 215–224. DOI: 10.1109/ICDE.2001.914830.

[5] F. Manganello, E. Rasca, A. Villa, A. Maddalena, and G. Boccuzzi, "Machine learning for post-diploma educational and career guidance: A scoping review in ai-driven decision support systems," *Front. Educ.*, vol. 10, p. 1 578 979, 2025. DOI: 10.3389/feduc.2025.1578979.

[6] A. Namoun and A. Alshanqiti, "Predicting student performance using data mining and learning analytics techniques: A systematic literature review," *Appl. Sci.*, vol. 11, no. 1, p. 237, 2020. DOI: 10.3390/app11010237.

[7] E. Kurilovas, I. Zilinskiene, and V. Dagiene, "Recommending suitable learning scenarios according to learners' preferences: An improved swarm based approach," *Comput. Hum. Behav.*, vol. 30, pp. 550–557, 2014. DOI: 10.1016/j.chb.2013.06.036.

[8] X. Zhou, A. Sun, J. Zhang, and D. Lin, "The crowd in moocs: A study of learning patterns at scale," *Interact. Learn. Environ.*, vol. 33, no. 3, pp. 2136–2150, 2025. DOI: 10.1080/10494820.2024.2392175.

[9] D. Wang, Y. Bao, G. Yu, and G. Wang, "Using page classification and association rule mining for personalized recommendation in distance learning," in *Proceedings of the International Conference on Web-Based Learning (ICWL)*, 2002, pp. 363–374. DOI: 10.1007/3-540-45689-9_30.

[10] S. Gunasekara and M. Saarela, "Explainability in educational data mining and learning analytics: An umbrella review," in *Proceedings of the 17th International Conference on Educational Data Mining (EDM)*, 2024, pp. 887–892. DOI: 10.5281/zenodo.12729987.

[11] N. Rohani, B. Rohani, and A. Manataki, "Clicktree: A tree-based method for predicting math students' performance based on clickstream data," *J. Educ. Data Min.*, vol. 16, no. 2, pp. 32–57, 2024. DOI: 10.5281/zenodo.13627655.

[12] M. J. Zaki, "Spade: An efficient algorithm for mining frequent sequences," *Mach. Learn.*, vol. 42, no. 1, pp. 31–60, 2001. DOI: 10.1023/A:1007652502315.

[13] X. Yan, J. Han, and R. Afshar, "Clospan: Mining closed sequential patterns in large datasets," in *Proceedings of the SIAM International Conference on Data Mining (SDM)*, 2003, pp. 166–177. DOI: 10.1137/1.9781611972733.15.

[14] N. Itzhak, S. Jaroszewicz, and R. Moskovitch, "Continuously predicting the completion of time-interval-related patterns," in *Proceedings of the Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD)*, 2023, pp. 239–251. DOI: 10.1007/978-3-031-33374-3_19.

[15] P. Garcia, A. Amandi, S. Schiaffino, and M. Campo, "Evaluating bayesian networks' precision for detecting students' learning styles," *Comput. & Educ.*, vol. 49, no. 3, pp. 794–808, 2007. DOI: 10.1016/j.compedu.2005.11.015.

[16] S. Graf, Kinshuk, and T.-W. C. Liu, "Identifying learning styles in learning management systems by using indications from students' behaviour," *J. Educ. Technol. & Soc.*, vol. 12, no. 2, pp. 123–137, 2009. [Online]. Available: https://www.jstor.org/stable/jeductechsoci.12.2.123.

[17] L. Hui, Y. C. Chen, and J. T. Y. Weng, "Incremental mining of temporal patterns in interval-based database," *Knowl. Inf. Syst.*, vol. 46, no. 2, pp. 423–448, 2016. DOI: 10.1007/s10115-015-0833-9.

[18] J. Zhang, J. Qiang, and C. Zhou, Eds., *New Horizons in Web Search, Web Data Mining, and Web-Based Applications*. MDPI Books, 2024. DOI: 10.3390/books978-3-7258-1656-5.

[19] M. Lee, A. Frenk, E. Worden, et al., "Investigating the robustness of knowledge tracing models in the presence of student concept drift," *arXiv preprint arXiv:2511.00704*, 2025. arXiv: 2511.00704 [cs.LG].

[20] O. Kassak, M. Kompan, and M. Bielikova, "Behavioral pattern mining for online time personalization," in *Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization (UMAP)*, 2017, pp. 325–326. DOI: 10.1145/3079628.3079640.

[21] Y.-C. Chen, W.-C. Peng, and S.-Y. Lee, "Mining temporal patterns in time interval-based data," *IEEE Trans. on Knowl. Data Eng.*, vol. 27, no. 12, pp. 3318–3331, 2015. DOI: 10.1109/TKDE.2015.2453952.

[22] H. Ngo, K. Vo, and T. Nguyen, "Personalized learning path recommendations: Fusing knowledge graph embedding, sequence mining, and collaborative filtering," in *Proceedings of the IEEE International Conference on Big Data (BigData)*, 2024, pp. 8145–8153. DOI: 10.1109/BigData62323.2024.10825001.

[23] C. Troussas, K. Chrysafiadi, and M. Virvou, "An intelligent adaptive fuzzy-based inference system for computer-assisted language learning," *Expert Syst. with Appl.*, vol. 127, pp. 85–96, 2019. DOI: 10.1016/j.eswa.2019.02.040.

[24] A. S. Al-Hegami and M. S. Kaity, "An ontology framework based on web usage mining," *Int. J. Appl. Inf. Syst.*, vol. 6, no. 9, pp. 28–35, Mar. 2014.

[25] S. Paulakis, C. Lampos, M. Eirinaki, and M. Vazirgiannis, "Sewep: A web mining system supporting semantic personalization," in *Proceedings of the European Conference on Principles of Data Mining and Knowledge Discovery (PKDD)*, 2004, pp. 556–558. DOI: 10.1007/978-3-540-30116-5_60.

[26] D. Gostautaite and R. Butkiene, "Recommender systems and recommendation dashboards for learning personalisation," in *Proceedings of the International Technology, Education and Development Conference (INTED)*, 2019, pp. 3591–3600. [Online]. Available: https://etalpykla.vilniustech.lt/handle/123456789/126317.

[27] D. J. Mustafa and S. R. M. Zeebaree, "Optimizing online educational experiences through semantic recommender systems: A review," *Asian J. Res. Comput. Sci.*, vol. 18, no. 10, pp. 211–226, 2025. DOI: 10.9734/ajrcos/2025/v18i10533.

[28] B. Kitchenham, "Procedures for performing systematic reviews," Keele University, Technical Report TR/SE-0401, 2004.

[29] B. J. Shea, B. C. Reeves, G. Wells, et al., "Amstar 2: A critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both," *BMJ*, vol. 358, j4008, 2017. DOI: 10.1136/bmj.j4008.

[30] M. J. Page, J. E. McKenzie, P. M. Bossuyt, et al., "The prisma 2020 statement: An updated guideline for reporting systematic reviews," *BMJ*, vol. 372, n71, 2021. DOI: 10.1136/bmj.n71.

[31] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Proceedings of the 31st International Conference on Neural Information Processing Systems (NeurIPS)*, 2017, pp. 4768–4777.

[32] M. T. Ribeiro, S. Singh, and C. Guestrin, "Why should i trust you?: Explaining the predictions of any classifier," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, 2016, pp. 1135–1144.

[33] E. Yafi, A. S. Al-Hegami, M. A. Alam, and R. Biswas, "Incremental mining of shocking association patterns," in *Proceedings of the 5th International Conference on Data Mining (ICDM 2009)*, January 28–30, 2009, Dubai, UAE, Jan. 2009.

[34] A. A. M. Mustafa, A. A. G. Al-Khulidi, and A. S. Al-Hegami, "A proposed incremental distance based clustering of interesting patterns for big data," *Sana'a Univ. J. Appl. Sci. Technol.*, vol. 4, no. 4, pp. 1938–1949, Apr. 2026. DOI: 10.59628/jast.v4i4.2271.

[35] Z. Lee, T. Lindgren, and P. Papapetrou, "Z-miner: An efficient method for mining frequent arrangements of event intervals," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, 2020, pp. 524–534. DOI: 10.1145/3394486.3403079.

[36] X. Li, Z. Zhang, and H. Wang, "Mining temporal characteristics of behaviors from interval events in e-learning," *Inf. Sci.*, vol. 447, pp. 169–185, 2018. DOI: 10.1016/j.ins.2018.01.048.

[37] O. Kassak, M. Kompan, and M. Bielikova, "Personalized recommendation in online systems using sequential pattern mining," *J. Intell. Inf. Syst.*, vol. 47, pp. 251–276, 2016. DOI: 10.1007/s10844-016-0410-5.

[38] M. R. Islam, A. M. Nitu, M. A. Marjan, et al., "Enhancing tertiary students' programming skills with an explainable educational data mining approach," *PLoS ONE*, vol. 19, no. 9, e0307536, 2024. DOI: 10.1371/journal.pone.0307536.

[39] S. Gunasekara and M. Saarela, "Explainable ai in education: Techniques and qualitative assessment," *Comput. & Educ. Artif. Intell.*, vol. 6, p. 100 234, 2025. DOI: 10.1016/j.caeai.2025.100234.

[40] O. Harel and R. Moskovitch, "Tirpclo: Efficient and complete mining of time intervals-related patterns," *Data Min. Knowl. Discov.*, vol. 37, no. 5, pp. 1806–1857, 2023. DOI: 10.1007/s10618-023-00935-5.

[41] M. Salehi and I. N. K. A. K. S. S. A. K., "Personalized recommendation of learning material using sequential pattern mining and attribute based collaborative filtering," *Educ. Inf. Technol.*, vol. 19, pp. 713–735, 2014. DOI: 10.1007/s10639-013-9254-2.

[42] A. S. Al-Hegami and H. Al-Omaisi, "Data mining techniques for mining query logs in web search engines," *IJCSN Int. J. Comput. Sci. Netw.*, vol. 6, no. 2, pp. 237–253, Apr. 2017. [Online]. Available: https://hal.science/hal-02276528.

[43] A. S. Al-Hegami, N. S. A. Al-Hddad, and H. A. Alsaeedi, "A web mining approach for evaluation of quality assurance at university of science and technology," in *Proceedings of the 4th Scientific Conference on Artificial Intelligence Technology and Quality Assurance in Higher Education*, Presented at the 4th Scientific Conference on Artificial Intelligence Technology and Quality Assurance in Higher Education, Sana'a, Yemen, Dec. 2023.