DOI 10.54596/2958-0048-2026-2-278-294
UDC 004.8
IRSTI 28.23.15

AN ANALYTICAL REVIEW OF MODERN ADAPTIVE METHODS FOR STREAM DATA PROCESSING IN REAL-TIME SYSTEMS

A. Jumagaliyeva[1*], M. Kaldarova[1], R. Ismailova[2], E. Abdykerimova[3], A. Turkmenbayev[3]
[1]Astana International University, Astana, Kazakhstan*
[2]Kyrgyz-Turkish Manas University, Bishkek, Kyrgyzstan
[3]Caspian State University of Technology and Engineering named after Sh. Yessenov, Aktau, Kazakhstan
[*]Corresponding author:mirakaldarova.zh@gmail.com

Abstract

Real-time systems process continuous, non-stationary data streams while adhering to strict latency constraints and employ adaptive learning mechanisms to ensure a stable level of prediction reliability. Current methods for achieving this goal typically prioritize either model accuracy or computational scalability over the integration of an adaptive learning method into a streaming environment. The objective of this study is to systematically review state-of-the-art adaptive learning methods for real-time streaming data processing using a multi-step methodology including bibliometric analysis, a systematic literature review, and a structured comparative synthesis. A total of 58 studies were analyzed to identify patterns in adaptive capabilities, architectural integration approaches, and overall system performance evaluation. The results revealed a consistent tradeoff between adaptivity and deterministic latency, as well as a lack of cross-layer coordination and performance measurement. Based on the analytical synthesis of the reviewed literature, a conceptual cross-layer analytical framework supporting the integration of adaptive learning and distributed streaming systems is proposed. Methodological recommendations for the design of adaptive systems demonstrating a high level of performance while maintaining stable operation in dynamic, non-stationary environments are presented.

Keywords: adaptive machine learning, stream processing, real-time systems, time latency, stability, interlayer system.

НАҚТЫ УАҚЫТ ЖҮЙЕЛЕРІНДЕ АҒЫНДЫ ДЕРЕКТЕРДІ ӨҢДЕУДІҢ ЗАМАНАУИ БЕЙІМДЕЛГІШ ӘДІСТЕРІНЕ АНАЛИТИКАЛЫҚ ШОЛУ

Джумагалиева А.[1*], Қалдарова М.[1], Исмаилова Р.[2], Абдыкеримова Э.[3], Туркменбаев А.[3]
[1]Astana International University, Астана, Қазақстан*
[2]Қырғыз-Түрік Манас университеті, Бішкек, Қырғызстан
[3]Ш. Есенов атындағы Каспий технологиялар және инжиниринг университеті, Ақтау, Қазақстан
[*]Corresponding author: mirakaldarova.zh@gmail.com

Аңдатпа

Нақты уақыт жүйелері қатаң кідіріс шектеулерін сақтай отырып, үздіксіз, стационарлық емес деректер ағындарын өңдейді және болжам сенімділігінің тұрақты деңгейін қамтамасыз ету үшін бейімделгіш оқыту механизмдерін қолданады. Бұл мақсатқа жетудің қазіргі әдістері әдетте адаптивті оқыту әдісін ағындық ортаға интеграциялаудан гөрі модель дәлдігіне немесе есептеу масштабталуына басымдық береді. Бұл зерттеудің мақсаты - библиометриялық талдау, жүйелі әдебиетке шолу және құрылымдық салыстырмалы синтезді қамтитын көп сатылы әдіснаманы қолдана отырып, нақты уақыт режимінде деректерді ағынмен өңдеудің заманауи бейімделгіш оқыту әдістерін жүйелі түрде шолу. Бейімделгіш мүмкіндіктердегі, архитектуралық интеграция тәсілдеріндегі және жалпы жүйенің өнімділігін бағалаудағы заңдылықтарды анықтау үшін барлығы 58 зерттеу талданды. Нәтижелер бейімделушілік пен детерминирленген кідіріс арасындағы тұрақты ымыраға келуді, сондай-ақ көп деңгейлі үйлестіру мен өнімділікті өлшеудің болмауын көрсетті. Осы нәтижелерге сүйене отырып, бейімделгіш оқыту мен таратылған ағындық жүйелерді интеграциялауды қолдайтын жаңа құрылым ұсынылады. Динамикалық, стационарлық емес ортада тұрақты жұмысты сақтай отырып, жоғары өнімділік деңгейін көрсететін бейімделгіш жүйелерді жобалау бойынша әдіснамалық ұсыныстар ұсынылған.

Кілт сөздер: бейімделгіш машиналық оқыту, ағынды өңдеу, нақты уақыт жүйелері, уақыт кідірісі, тұрақтылық, қабатаралық жүйе.

АНАЛИТИЧЕСКИЙ ОБЗОР СОВРЕМЕННЫХ АДАПТИВНЫХ МЕТОДОВ ПОТОКОВОЙ ОБРАБОТКИ ДАННЫХ В СИСТЕМАХ РЕАЛЬНОГО ВРЕМЕНИ

Джумагалиева А.[1*], Қалдарова М.[1], Исмаилова Р.[2], Абдыкеримова Э.[3], Туркменбаев А.[3]

[1]Astana International University, Астана, Казахстан*

[2]Кыргызско-Турецкий университет Манас, Бишкек, Кыргызстан

[3]Каспийский университет технологий и инжиниринга имени Ш. Есенова, Актау, Казахстан

[*]Корреспондент автор: mirakaldarova.zh@gmail.com

Аннотация

Системы реального времени обрабатывают непрерывные, нестационарные потоки данных, соблюдая при этом строгие ограничения по задержке, и используют механизмы адаптивного обучения для обеспечения стабильного уровня надежности прогнозирования. Современные методы достижения этой цели, как правило, отдают приоритет либо точности модели, либо вычислительной масштабируемости, а не интеграции метода адаптивного обучения в потоковую среду. Цель данного исследования – систематический обзор современных методов адаптивного обучения для обработки потоковых данных в реальном времени с использованием многоэтапной методологии, включающей библиометрический анализ, систематический обзор литературы и структурированный сравнительный синтез. Всего было проанализировано 58 исследований для выявления закономерностей в адаптивных возможностях, подходах к архитектурной интеграции и оценке производительности всей системы. Результаты показали постоянный компромисс между адаптивностью и детерминированной задержкой, а также отсутствие межслойной координации и измерения производительности. На основе аналитического синтеза рассмотренной литературы предлагается концептуальный межслойный аналитический фреймворк. Представлены методологические рекомендации по проектированию адаптивных систем, демонстрирующих высокий уровень производительности при сохранении стабильной работы в динамических, нестационарных средах.

Ключевые слова: адаптивное машинное обучение, потоковая обработка, системы реального времени, задержка времени, стабильность, межслойная система.

Introduction

Real-time sensing systems are increasing in their capability to sense and act without delay, while Cyber-Physical Systems and their associated distributed data acquisition processes are escalating the data acquisition rate through unbounded streams of data that must be processed in real-time to enable fast decision making. In general, modern real-time systems exhibit multiple characteristics: they generally have strict latencies; the data streams are heterogeneous and can be very high-frequency; and there can be dynamic changes to the data streams. Consequently, stream processing system designs have also evolved from a static batch processing model to an adaptive, real-time modular architecture that incorporates continuous learning and dynamic resource management.

Many traditional stream analytics systems are based on machine-learning models that have been trained offline and have fixed-window mechanisms. Some stream analytics approaches are filtering at a statistical level, extracting time/frequency-based features, or employing supervised machine-learning methods, such as support vector machines (SVM), random forest, k-nearest neighbors k-NN, multi-layer perceptron (MLP), convolutional neural network (CNN), or long short-term memory (LSTM)). While these types of systems have demonstrated a high level of accuracy within controlled environments, they typically perform as static inference systems when subjected to real-time processing. As a consequence, these systems were designed without considering incremental updates or adaptations due to drift in the distribution of data generated by a stream processing system.

To enable large-scale event processing while also providing the required properties of scalability, parallelism, and throughput, a number of distributed event-processing systems have been developed, such as Apache Storm, Apache Flink, and Spark Streaming. Events processed through these systems use a general-purpose data-flow model that does not provide direct support for adaptive learning. Modeling efforts that seek to update models, address drift, and optimize use of resources are provided through external systems rather than as part of the distributed event-processing environment.

Concept drift, which occurs when the statistical properties of the data change over time, represents one of the most significant challenges encountered when working with real-time streaming data. The concept of drift can be attributed to a wide variety of factors, including environmental changes, configuration changes to an event-processing system, changes in user behavior, and sensor degradation. To address the need for non-stationary data, many adaptive methodologies have been developed, including (i) online/incremental learning methods based on stochastic gradient descent (SGD), Hoeffding trees, or adaptive ensembles; and (ii) drift-detection methodologies such as DDM and ADWIN. A disadvantage of these adaptive methods is that they result in an increase in the computational cost, memory footprint, or latency of a deterministic real-time system.

A use case that clarifies the necessity for scalable adaptations needed to support continuous real-time streaming is the processing of biomedical signals. Continuous-processing systems monitoring ECG, EEG, EMG, and wearable devices must provide low latency, resistance to noise, and sensitivity to gradual/class-based shifts in the subject's physiological state. A delay in adaptation within a continuous-processing system performing biomedical streaming could reduce the reliability of that system. Thus, biomedical streaming clearly illustrates the need for high-performance, unified, adaptive architectures.

Most of the current distribution and adaptive algorithm research focuses primarily on maximising accuracy of prediction or processing speed. Relatively few frameworks exist to support adaptive stream-processing models that jointly provide (i) awareness of drift, (ii) incrementally update system state, (iii) maintain a stable level of latency, and (iv) adopt scalable architectures. As a result, there does not exist a comprehensive analytical synthesis of adaptive stream-processing approaches under real-time constraints.

Research Questions

In this study, the authors investigate:

1) How to combine adaptive online learning with real-time streaming architecture while minimizing the effect of non-deterministic latency?

2) What are the trade-offs in terms of performance, scalability, and adaptability in contemporary streaming systems?

3) How can drift detection, incremental updating, and resource-aware optimization function together within an adaptive framework?

Objectives

This paper has three objectives:

1) Present a detailed analytical review of adaptive streaming processing techniques in real-time environments.

2) Establish a systematic taxonomy of the existing methods, based upon levels of architecture, algorithm, and computational characteristics.

3) Identify potential areas of research and suggest a synthesised theoretical framework that combines adaptive learning and high-performance streaming techniques.

The primary analytical contribution of this review derives from the comparative cross-study synthesis of adaptive learning mechanisms, distributed streaming architectures, and real-time performance constraints.

The analytical review presents a unified analytical framework that communicates the interactions among adaptability, latency stability, and computational efficiency; a structured taxonomy of current adaptive streaming methods; and a side-by-side comparison of those methods according to their latency, degree of scalability, and resource usage. The authors' primary focus was to identify key trade-offs between adaptability and deterministic real-time performance, and to develop an analytical foundation for integrating adaptive learning into high-performance streaming architectures.

Literature Overview

Research on real-time adaptive stream data processing exists within many different research paradigms that frequently have varying definitions of real-time, many studies evaluate their systems with some other assumptions for deployment and report performance using a variety of computational metrics which makes it difficult for others to compare amongst the studies [1-3]. As it relates to adaptive stream processing, once can say that “real-time” in general, relates to analysing an unbounded but continuous flow of data under strict constraints based on time. However, the meaning of latency tolerance, critirism of determinism, and frequency of adjustments varies greatly across the literature. As a result, success rates of adaptive methods are usually more reliant upon implicit architectural issues rather than any soundly defined overarching criteria.

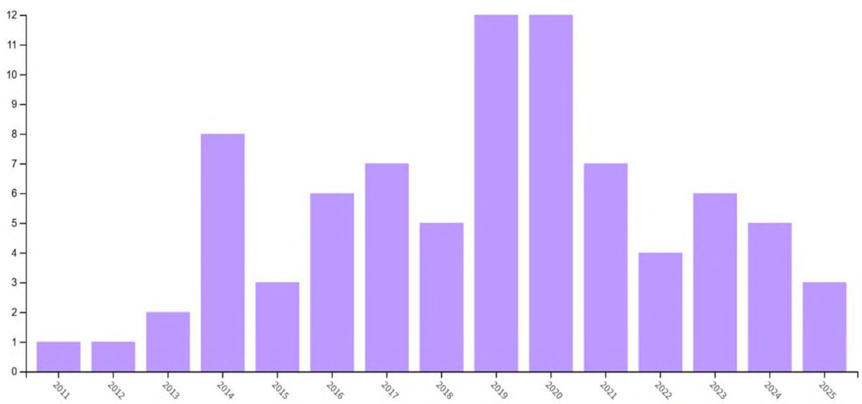

Figure 1. Number of articles published per year for analysis indexed in the Web of Science database

In the literature there is notably continued growth of published literature that is indexed within the Web of Science Core Collection in the data years of 2011-2025 around adaptive stream data processing in real-time systems as shown in Fig. 1. The literature shows that there was very little published in the early 2010s with a significant growth curb of literature in the time-frame of 2014-2015. There is a significant increase in growth rate during the period of 2018-2021, showing the highest levels of output per year [4].

Distributed streaming frameworks have grown in popularity alongside newly developed approaches for online and incremental learning. Also driving demand is the growing interest in using IoT, cyber-physical systems, and healthcare data for real-time analytics. The stability seen after 2021 is attributed to the maturation of the field, rather than a decrease in research activity, and thus adaptive real-time stream processing is recognized as an established and stable area of research.

Contemporary real-time processing systems must process a flow of heterogeneous data as it arrives continuously; the underlying distribution of the data may change randomly over time and they need to have deterministic responses to incoming events while using few computational resources. Therefore, unlike batch processing systems, streaming architectures are not able to rely on having access to all of the data prior to processing or that the distributions of the data will be stable [5-6]. Streaming architectures must be able to update their models incrementally and to provide high performance without having access to all of the data from which to retrain their models. Thus, streaming architectures have adapted by employing incremental learning, drift-aware adaptation, and resource-aware scheduling mechanisms.

Evaluating real-time efficiency poses another challenge. Most studies evaluate only average latencies without taking into account ingestion overhead, window processing latencies, broker latencies, or synchronization costs; conversely, some studies measure throughput based solely upon a fixed input rate without accounting for the impact of bursty workloads on performance [6]. Memory usage and state complexity are typically ignored as well. In addition, in an adaptive real-time stream processing system, the frequency of model updates, drift detection overhead, and the cost of replacing models further complicate evaluations of platform performance. Consequently, real-time efficiency cannot be evaluated using only a single dimension; instead, it should be considered using at least five different dimensions: latency, throughput, memory usage, adaptation cost, and scalability [7].

The literature on adaptive real-time stream processing is still fragmented. The research on machine learning tends to focus on online optimization, incremental trees, ensembles, and drift detection. The research on streaming systems tends to focus on windowing, state management, backpressure, fault tolerance, and scalability within distributed processing frameworks such as Flink and Spark. The research related to high performance typically focuses on the use of GPUs or edge computing for performance acceleration. These different directions of research are often analyzed in isolation resulting in only partial optimizations of either adaptability or computational performance, but not of both concurrently.

A recurring pattern observed across the reviewed studies is that improvements in predictive adaptability are frequently associated with increased computational overhead, synchronization complexity, or latency instability. Conversely, studies emphasizing deterministic throughput and scalability often rely on static or weakly adaptive learning mechanisms. This demonstrates that adaptability and real-time computational stability are still commonly optimized independently rather than through coordinated cross-layer system design.

Figure 2. Web of Science Core Collection, Clarivate Analytics

The subject category distribution shown in Fig. 2 helps clarify the structural position of research on adaptive stream data processing. Most of the publications listed are in Computer Science Theory and Methods, Computer Science Information Systems, and Computer Science Artificial Intelligence, indicating that adaptive stream data processing is primarily based on algorithmic development and systems-level designs.

A large proportion of the research comes from Electrical and Electronic Engineering and is indicative of the hardware and infrastructure aspects of real-time systems. The interdisciplinary distribution of research further indicates that adaptive stream processing is located at the intersection of learning algorithms, distributed systems, and applied analytics [8].

As such, adaptive stream processing is not a single-method activity but in fact a cross-layer domain that must integrate algorithms, streaming infrastructure, and hardware-aware optimization for the ultimate goal of being able to process streams of data in real time.

It also poses some challenges with respect to evaluation methodology; the majority of the literature evaluates adaptive processing using simulated data streams via the replay of previously recorded data or through the use of an artificial sliding window without the conduct of long-term tests, and there are significantly different methods for defining data drift that are simulated and may or may not reflect actual gradual shifts in real-world applications. There are also considerably different configurations of hardware, ingestion pipelines and external systems that are used; therefore, comparisons are difficult. The lack of standardized definitions for end-to-end latency or adaptation costs makes it impossible to determine if an improved result is the product of improvements to the algorithm being optimized or merely the result of changes to system configurations.

Experimental configurations, drift simulation methodologies, hardware environments, and latency reporting practices differ substantially across the reviewed studies. As a result, many reported performance improvements remain highly context-dependent and cannot yet be interpreted as universally transferable across heterogeneous real-time streaming environments.

One domain that illustrates these issues is biomedical signal processing, because there are stringent latency requirements along with significant non-stationarity and noise in the data. Therefore, adaptation must occur within the prescribed bounded response times and maintain diagnosis reliability. Although biomedical signal processing reflects only one research domain, it illustrates how adaptation, latency control and resource management interact under stringent constraints.

For this reason, adaptive stream processing should be considered a cross-layer systems issue, rather than merely an algorithmic one. In study [9], adaptation mechanisms, architectural integrative strategies and system-level performance indicators are reviewed to identify the shared tradeoffs and the existing research gaps leading to an integrated framework for these three research domains in the subsequent sections.

The comparison of adaptive stream processing research is meaningful only when methods are assessed in the context of architecture integration and system-level performance. The papers [10-11] reflected the adaptation mechanism, streaming integration and computation performance are assessed as a function of three dimensions in this Section.

The majority of papers investigate predictive accuracy with simulated streams. Deep neural networks have been found to be accurate when used within the biomedical, IoT and cyber-physical domains [12]. However, in most cases they have been evaluated in the context of simulated streaming with sliding window replay, and their long-term stability under a variety of load conditions is often not evaluated.

Paper [13] discussed distributed frameworks, such as Flink, Spark, Kafka and report on scalability, throughput, and fault tolerance, however many have static models and retrain outside of the streaming loop. Therefore, just because an infrastructure can scale, does not mean it will also be adaptive.

The most recent studies [14-15] presented new methodologies that incorporate self-learning, online learning, incrementally growing trees, ensemble classifiers, and drifting detection capabilities, like ADWIN and DDM. These processes make an adaptive system more robust to distribution shifts, but they do not frequently report on the cost of adaptation (e.g., update latency, synchronization overhead, or memory growth). Trade-offs between being adaptive and deterministic in performance are regularly under-evaluated.

The use of GPUs, multi-core processors, and deploying edge can decrease latency, and increase throughput, but they have typically only been evaluated with static models. The interaction between continuous adaptation and hardware acceleration continues to be understudied. The comparative synthesis of the reviewed studies revealed three recurring structural patterns.

First, predictive accuracy is frequently evaluated independently from full end-to-end system performance;

Second, the presence of scalable distributed infrastructure does not necessarily guarantee effective adaptive capability or coordinated real-time optimization;

Third, adaptive resource management and system-level architectural coordination are rarely implemented jointly.

All of these patterns demonstrate the lack of a systematic approach to facilitating systems integration at the current functional level of integration (integration of multiple computers, servers, and various types of software).

The 2022-2026 representative studies listed within Table 1 provide substantial evidence for these patterns not only due to their different scope of application but due to their varying methodology, as well as how these different methodologies lead to results that can be verified. Ultimately, the use of different methodologies and the many documented instances of different results not only provide evidence for the accuracy or inaccuracy of these representative studies but also for the gaps that exist in the field of research and development and for the multiple types of trade-offs that may exist among the separate studies.

Table 1. Comparative overview of recent research on adaptive stream processing (2022-2026)

Application Domain	Methodological approach	Streaming data exchange methodology	Main results
Biomedical monitoring domains	CNN and LSTM Deep learning models	Scenario-based planning strategy	High-fidelity, low-latency analytics
IoT in Industry	Transformer models	Spark Streaming style processing	Stable performance throughout steady state
Analytics of finance	Online learning (SGD)	Streaming continuous strategy	Resistant to slow changes
Cybersecurity cases	Drift-aware ensembles	Distribution strategy	Rapid management of sudden changes in distribution
Healthcare integrated devices	Incremental models	Low-latency real time pipeline	Reduced training overhead
Infrastructural Smart domain	Hoeffding Tree	Event-driven architecture strategy	Flexible learning paths in real time
Analytics of Edge domain	GPU-accelerated models	Mixed edge and cloud computing infrastructure strategy	Reducing synchronization delays

Table 1 supports the analytical finding that contemporary adaptive stream processing studies optimize predictive performance, scalability, and adaptive capability unevenly across different architectural layers rather than through fully coordinated system-level integration.

Current limitations in structural limitations and systemic weaknesses within current biomedical streaming processing evidence structures demonstrate relevant barriers to providing robust and reproducible biomedical streaming systems that are effectively implemented and used in practice. Predictive modelling, the development of streaming infrastructure and adaptive learning have progressed quite significantly through research and through the analysis of multiple studies relating to each area; however, some key structural limitations will very likely inhibit the robustness and reproducibility of biomedical streaming systems.

These structural limitations consist of significant disconnects between the adaptive learning capabilities of most current biomedical streaming systems and the underlying architecture used to support streaming processing; as many current systems are designed for adaptive learning under the flawed assumption that the systems will be used in a static deployment when attempting to simulate a continuous streaming system, thus limiting the systems' ability to withstand gradual distributional shifts. While adaptive approaches increase the adaptability of biomedical streaming systems, they often incur greater computational overhead and introduce latency instability. Distributed frameworks, while improving the scalability of streaming processing, remove the capability of adaptive learning to use the core processing logic, thereby degrading the performance of streaming systems. Hardware acceleration improves inference times; however, it is presupposed that loaded data will remain stable over time and the synchronisation cost incurred between the hardware and the software part of the streaming processing system has not been considered.

Also, methodological heterogeneity is a significant limitation in comparing the performance of existing biomedical streaming processing systems, as the methodologies used across studies generally provide only mean latency and throughput metric reports and do not typically provide reports on tail latency, jitter, long-term stability or end-to-end performance profiling. Table 2 presents a synthesis of common barriers identified across the studies.

Table 2. Identified limitations and possible directions of solution

Identified limitation	Systemic cause	Potential research direction
Static model assumptions	Lack of built-in adaptation	Drift-aware online learning integration
Latency instability	Update and sync overhead	Resource-aware adaptive scheduling
Architecture–model separation	Disjoint system design	Cross-layer unified frameworks
Incomplete performance evaluation	Focus on average metrics	Full latency distribution analysis
Adaptation–efficiency trade-off	Accuracy-centered evaluation	Joint optimization strategies
Hardware–adaptation mismatch	Static workload assumptions	Adaptive GPU/edge coordination
Experimental inconsistency	Heterogeneous setups	Standardized evaluation protocols
Missing end-to-end profiling	Pipeline cost ignored	Full system-level performance modeling

Table 2 supports the analytical conclusion that many current limitations in adaptive stream processing originate from structural fragmentation between adaptive learning mechanisms, streaming architectures, and system-level performance management.

Currently, biomedical research using streaming data is insufficiently coordinated across all relevant areas, including algorithm development, architecture design, and hardware optimization. Identifying ways to bridge the existing gaps between these areas will require a multi-layered collaborative approach that considers training adaptation, streaming data integration, and real-time performance limitations across multiple parameters.

Materials and methodology

The research methodology in this study is a process of analytical review that includes three components: bibliometric analysis, systematic selection and filtering of studies for inclusion in the analysis, and a structured synthesis of methodological and architectural models. The overall goal of the methodology is to provide a transparent, reproducible, and comparable way of synthesizing research from multiple disciplines; while allowing, at the same time, for flexibility in the analysis needed for interdisciplinary synthesis.

There are three parts that make up the review process. The first part of the process is bibliometric landscape analysis. The second part is the systematic selection and filtering of studies. The third part of the method is the structured analysis of the methodology and structure of the included studies to produce a synthesis of the findings. We used the Web of Science (WOS) Core Collection database to conduct the Bibliometric analysis. Publications were retrieved based on a set of defined keywords relevant to Adaptive Learning, Stream Data Processing, Online Learning, Concept Drift, Distributed Stream Processing, and Real-

Time/Embedded Systems. Only English language articles and conference proceedings published in peer-reviewed journals that appeared between 2011 and 2025 are included.

To improve methodological transparency and reproducibility, the literature search strategy was constructed using controlled combinations of keywords associated with adaptive learning, online stream processing, concept drift, distributed architectures, and real-time systems. Additional verification of highly cited publications was performed through indexed scientific databases to reduce the possibility of excluding influential methodological studies relevant to adaptive streaming research.

The generalized search query included combinations of the following terms: (“adaptive learning” OR “online learning” OR “incremental learning”) AND (“stream processing” OR “data streams”) AND (“real-time systems” OR “real-time analytics”) AND (“concept drift” OR “drift detection”). The search syntax was iteratively refined to improve thematic relevance and reduce overlap between unrelated studies.

Main Category	Proportion of Total Publications	Top Subcategories	Proportion within Main Category
Computer Science	58.0%	Computer Science Theory & Methods	21.4%
		Computer Science Artificial Intelligence	18.7%
		Computer Science Information Systems	10.2%
		Computer Science Software Engineering	7.7%
Engineering	24.0%	Engineering Electrical & Electronic	11.3%
		Engineering Industrial	5.4%
		Automation & Control Systems	4.1%
		Engineering Multidisciplinary	3.2%
Mathematics & Statistics	8.5%	Applied Mathematics	4.9%
		Statistics & Probability	3.6%
Interdisciplinary & Applied Domains	9.5%	Robotics	3.1%
		Telecommunications	2.7%
		Biomedical Engineering	2.3%
		Smart Systems / IoT	1.4%

Figure 3. Disciplinary distribution of research into specific categories

According to the data presented in Fig. 3, scientific categories involved in the research landscape are organisationally represented according to the contribution made by each discipline to the development of adaptive stream processing. Based on this information, ASP is primarily associated with the field of computer science, in addition to a secondary relationship with engineering disciplines. Theoretical foundations for algorithmic modelling come from mathematics, while applied contextually validating examples are available in multi-disciplinary research areas, such as biomedical engineering and internet of things (IoT) systems.

This bibliometric map establishes the disciplinary basis for the adaptive stream processing field and the structural framework for methodological classifications used in comparative studies across research domains.

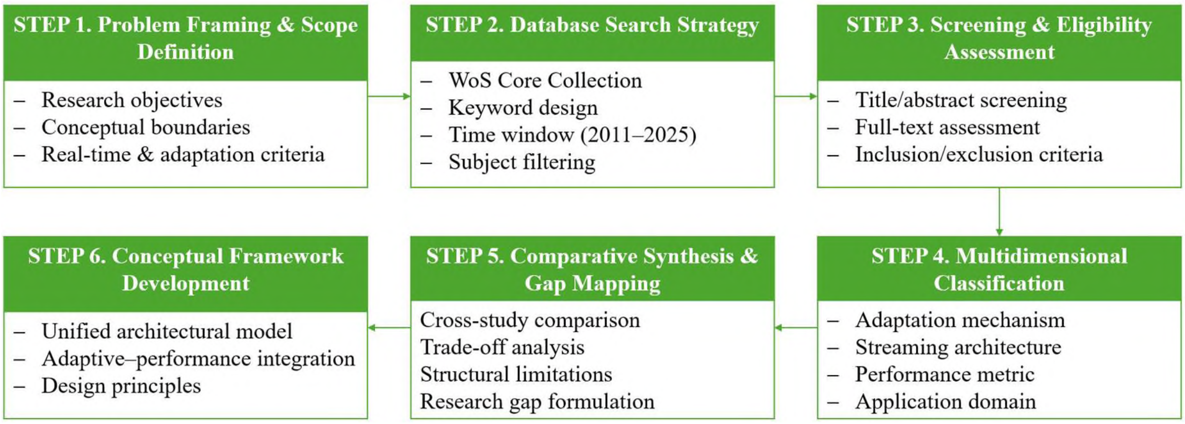

Figure 4. Structured analytical workflow of the review methodology

This is a visual representation of the structured analytical workflow used for conducting this review as it shown in Fig. 4. To begin with, the first stage of the review involves defining a problem framing and defining the objectives of the research and real-time research adaptation criteria. The second stage consists of performing a controlled keyword-based database search using the Web of Science Core Collection as a primary source, while also enforcing time-boundaries and filtering for subject matter of interest to this review. Third, following retrieval from the database, each study retrieved is subject to text review with respect to specific inclusion and exclusion criteria as identified in the search process.

Studies were included in the review if they satisfied several conditions: publication in peer-reviewed journals or indexed conference proceedings, relevance to adaptive or drift-aware stream processing, presence of a real-time or near real-time operational context, and availability of quantitative evaluation metrics such as latency, throughput, scalability, or adaptation overhead. Studies focused exclusively on static batch-processing systems, conceptual discussions without implementation details, or publications lacking quantitative validation were excluded from the final analytical synthesis.

After the final classification of all selected articles according to multiple analytical dimensions, including adaptation mechanisms, streaming architecture, performance criteria, and application domain, the selected studies were comparatively synthesized to identify structural deficiencies, recurring research gaps, and cross-layer integration patterns. This process additionally supported the development of a synthesized analytical framework for adaptive stream data processing in real-time systems derived from comparative cross-study analysis.

To improve analytical consistency, the selected studies were additionally evaluated according to methodological quality dimensions including reproducibility of experimental configuration, completeness of latency and throughput reporting, architecture-level integration, scalability validation, and the implementation of adaptive mechanisms. Publications containing incomplete methodological descriptions or limited system-level reporting were considered partially representative within the comparative synthesis.

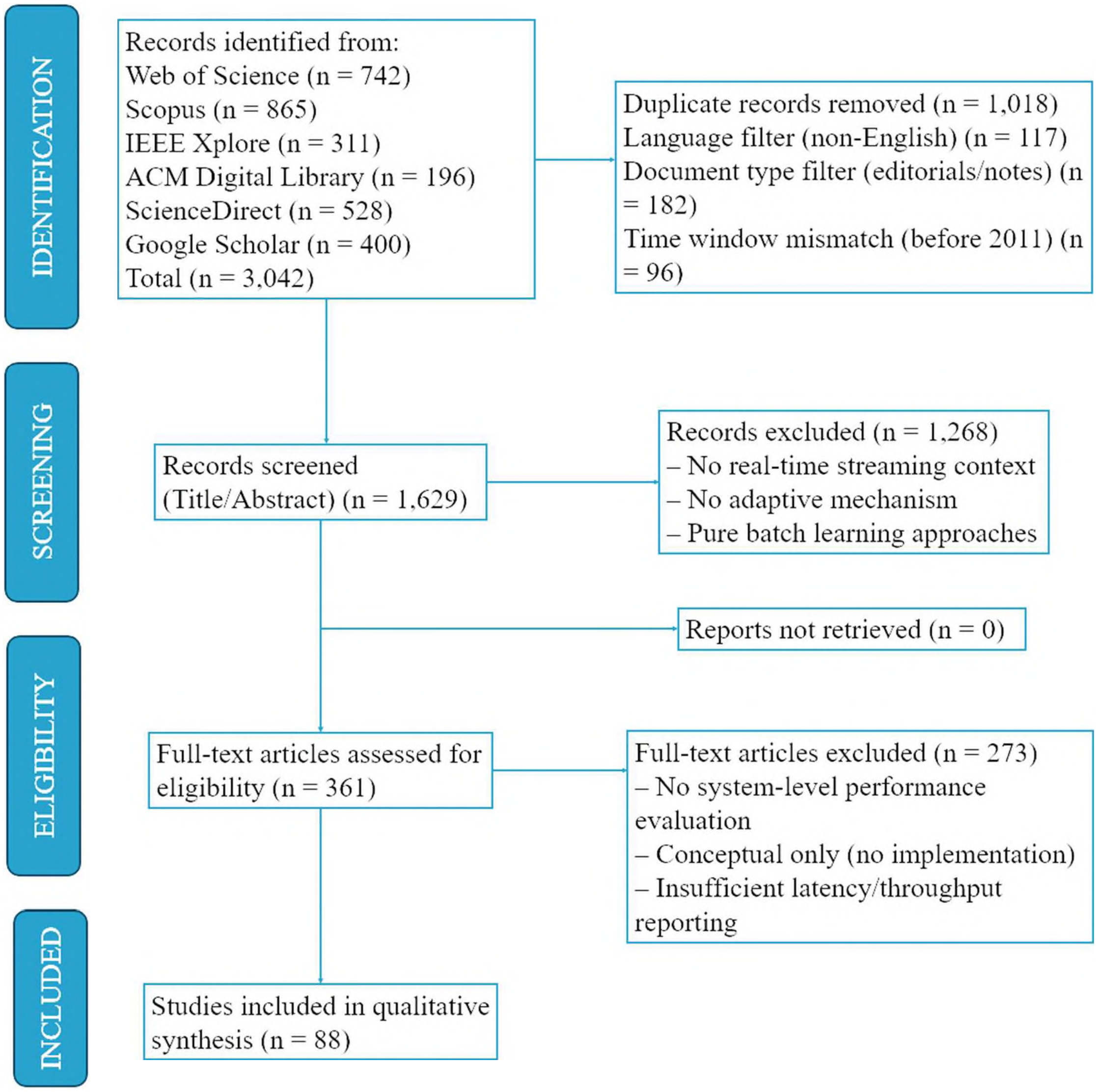

Figure 5. Study selection process based on PRISMA

Fig. 5 shows the PRISMA-based flowchart. It was used to select the studies for the adaptive streaming data processing research between 2011 and 2025. A total of 742 records were taken from the Web of Science database, using controlled keywords for adaptive learning, streaming data processing, concept drift, and real-time systems. After filtering these records by document type, language, date range, and subject area, there were 491 documents left to be reviewed for titles and abstracts. During the first stage of screening, 326 studies were identified as having no streaming context in real-time or had not used adaptive systems or only addressed batch operations; thus, those 326 studies were removed from further evaluation. Therefore, there are 165 articles remaining to examine for eligibility against the evaluation criteria related to the system including actual latency and what the throughput, scalability and adaptive performance. Following this process, 107 studies were removed due to no quantitative performance validation or implementation details, leaving 58 studies to be included within the final qualitative assessment and comparative analysis.

The PRISMA-based selection procedure ensured a transparent multi-stage filtering process. Duplicate and weakly relevant publications were removed during title and abstract screening, while eligibility assessment focused on adaptive capability, real-time deployment relevance, and the presence of quantitative system-level evaluation metrics. Studies lacking reproducible implementation details or comprehensive experimental validation were excluded to strengthen the methodological reliability of the final review corpus.

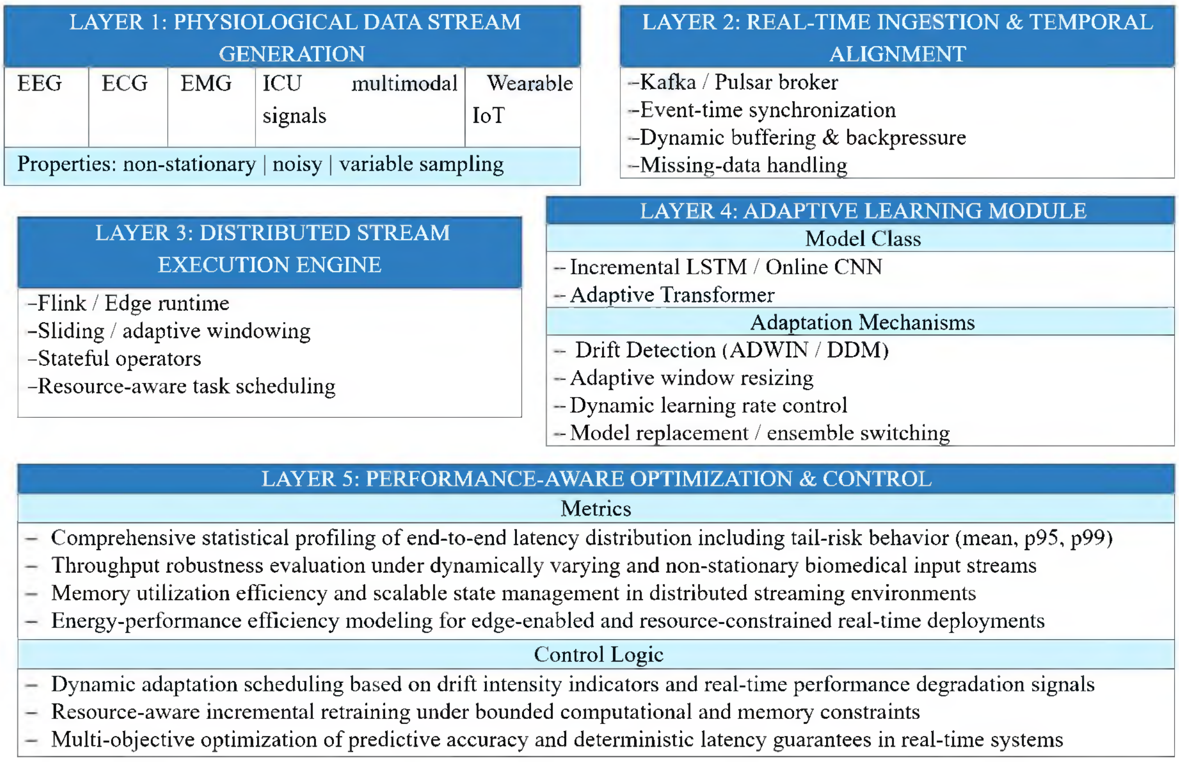

Figure 6. Multi-layer architecture for real-time processing of biomedical streaming data

Fig. 6 illustrates the conceptual analytical framework synthesized from the reviewed literature for cross-layer adaptive processing of biomedical signal streams in real-time environments. Data collected from various heterogeneous sources, such as EEGs, ECGs, EMGs, ICUs and wearable devices can be processed in a distributed fashion and learned from by incorporating embedded adaptive learning mechanisms into one continuous closed-loop system. Incremental deep models and drift detection methods are able to run directly within the streaming workflow, while the performance-aware control layer continuously sweeps through the latency of the processes; distribution of processed streams based upon throughput stability and memory use; and energy efficiency. Dynamic adaptation occurs in realtime within defined constraint boundaries to maintain consistent high-quality predictions without exceeding the deterministic deadlines for completion.

Results and Discussion

The structured analysis of 58 studies demonstrated that adaptive streaming processing research has evolved unevenly across different methodological and architectural directions during the period from 2011 to 2025. Biomedical applications are the most prominent and are followed by both industrial IoT and smart infrastructure to demonstrate the strong connection between non-stationary environments and the necessity for making real-time decisions. The comparative synthesis across the reviewed studies demonstrates that many adaptive stream processing approaches continue to optimize isolated system properties independently rather than implementing fully integrated cross-layer adaptive coordination.

These findings are further supported by conceptual analysis, indicating that adaptive models improve robustness under concept drift, while introducing additional computational overhead and latency variability. Among the reviewed studies, adaptive ensemble methods, incremental learning models, and drift-aware architectures represented the dominant methodological approaches. However, only a limited subset of publications reported comprehensive end-to-end system evaluation metrics including tail latency, jitter analysis, or long-term operational stability.

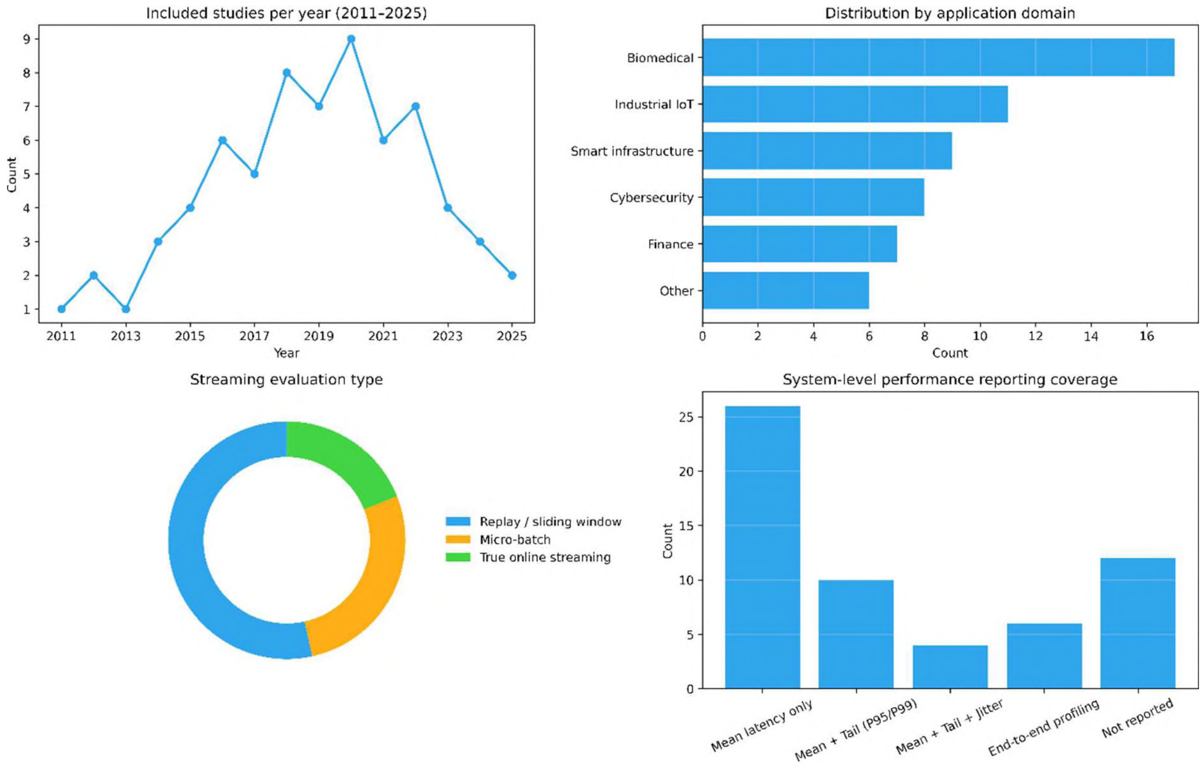

Figure 7. Results of the empirical distribution of studies found in this review

Fig. 7 illustrates how real-time processing has considerable focus from a conceptual standpoint, yet the evaluation methodologies applied across the reviewed studies remain highly heterogeneous. Empirical evaluations are performed predominantly through replay or sliding window simulation methods, whereas the number of actual online streaming installations is relatively low. Further, the most frequently reported metrics at the system level have been average latency and infrequent reports of tail latency or end-to-end profiling, suggesting that comprehensive reporting of unpredictable operational behaviour remains limited.

This analytical observation demonstrates that evaluation methodology standardization remains insufficient across adaptive real-time streaming research.

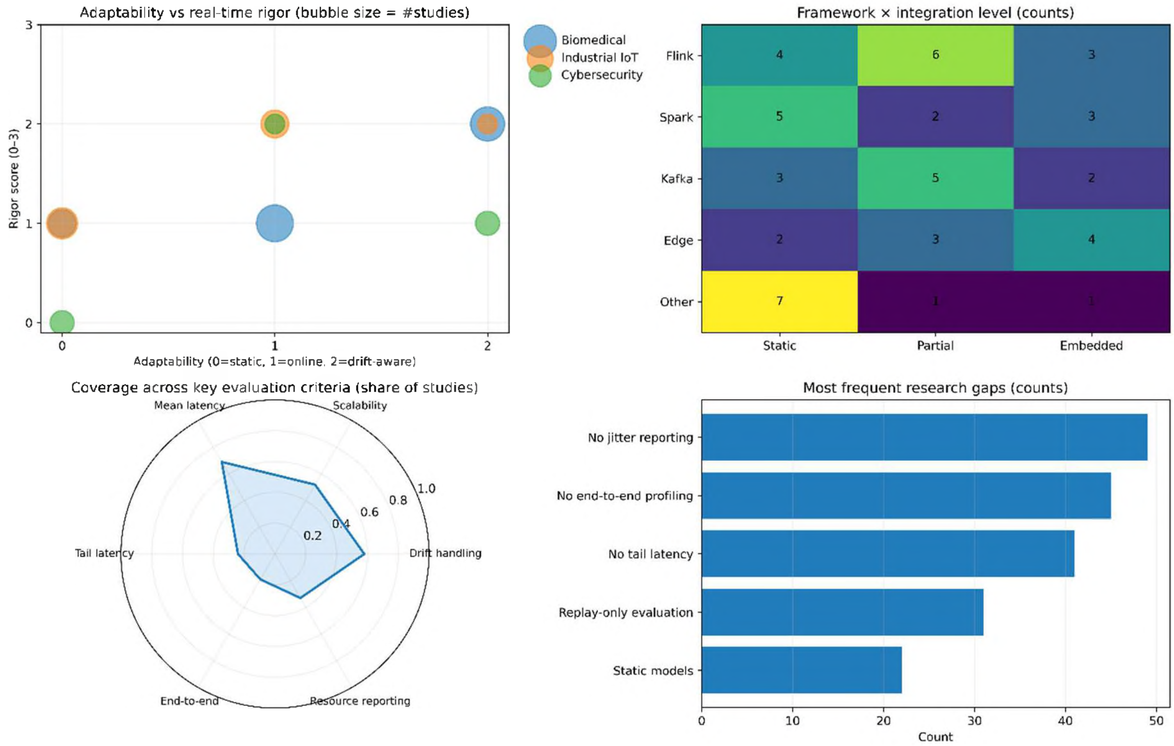

Figure 8. Results of structural constraints patterns

The analysis of the layers in Fig. 8 indicates that the trade-offs between rigorous structural evaluation and structural adaptability are manifested with the operational mechanisms of drift awareness and online adaptation providing greater robustness to distribution shifts as well as being more broadly reported in terms of latency distributions and resource accounting. Conversely, while there is a more consistent performance report associated with static models and partially adaptive systems, there is also increased fragmentation between innovations to algorithms and their validation against architecture.

In addition, based upon cross-layer integration matrices and rates of gap distribution shown in Figure 8, that there is no inherent embedded adaptability provided by the distribution framework in which scalability exists. As a result, the presence of static deployments continues to exist within scalable infrastructure environments, thereby maintaining significant research gaps related to tail latency analysis and end-to-end operational visibility. In summary, the findings of this research demonstrate that while the fields of adaptive streaming processing research have been growing across individual dimensions of adaptability, they have yet not been developed as fully coordinated systems. Fig. 8 supports the analytical finding that increasing adaptive capability is frequently associated with higher architectural complexity, synchronization overhead, and reduced determinism of real-time performance.

Despite substantial progress in adaptive stream analytics, the reviewed literature continues to demonstrate fragmentation between adaptive learning algorithms, distributed streaming infrastructures, and deterministic real-time performance management. Most contemporary approaches optimize isolated system dimensions independently, including predictive accuracy, throughput, scalability, or adaptation capability, while comparatively few studies propose fully coordinated cross-layer adaptive architectures suitable for robust deployment in highly dynamic operational environments.

Conclusion

The Meta-analysis examined 58 studies of on-line adaptation for stream processing in real-time systems and found that there is increasing evidence of evolution towards drift-aware and scalable architectures related to the stream processing systems. The comparative analytical synthesis additionally revealed persistent structural fragmentation between adaptive mechanisms, streaming infrastructures, and deterministic real-time performance evaluation approaches. Additionally, while predictive robustness has improved, the systematic reporting of performance characteristics such as latency distributions, cost of updating, or cost of going from ingestion to decision making have also been very limited. The analytical findings of the review highlight the importance of balancing adaptability and deterministic latency stability within real-time streaming systems. Future research should focus on lightweight adaptive mechanisms that minimize computational overhead while preserving model responsiveness under dynamic data conditions.

Key Analytical Findings of the Review

1) The comparative analysis of the reviewed studies demonstrated a consistent trade-off between predictive adaptability and deterministic latency stability in real-time streaming systems.

2) Cross-study synthesis revealed that current adaptive stream processing research remains fragmented across algorithmic, architectural, and infrastructure layers, limiting coordinated system-level optimization.

3) The analytical review identified that most existing studies evaluate average latency and throughput while insufficiently reporting tail latency, jitter, synchronization overhead, and adaptation cost.

4) The reviewed distributed streaming frameworks frequently provide scalability and fault tolerance but lack embedded mechanisms for coordinated adaptive learning and drift-aware optimization.

5) The synthesis of the reviewed literature indicates that unified cross-layer adaptive architectures remain one of the primary unresolved challenges for robust real-time intelligent streaming systems.

An overall conclusion from this study is that there is a strong need for unified cross-layer frameworks for adaptive learning, integration of architectures, and multi-dimensional real-time performance constraints to achieve better results with reliable real-time intelligent systems in biomedical, industrial, and cyber-physical domains. Future research should focus on developing standardized evaluation protocols and adaptive scheduling strategies that maintain latency stability during dynamic changes in the evolution of the data distribution.

References:

1. Alam M.A., Nabil A.R., Mintoo A.A., Islam A. Real-time analytics in streaming big data: techniques and applications // Journal of Science and Engineering Research-2024. -Vol. 1(01). -P. 104-122. DOI: https://doi.org/10.70008/jeser.v1i01.56.
2. Almeida A. et al. Time series big data: a survey on data stream frameworks, analysis and algorithms //Journal of Big Data. – 2023. – Т. 10. – №. 1. – С. 83. DOI: https://doi.org/10.1186/s40537-023-00760-1.
3. Marcu O.C., Bouvry P. Big data stream processing // Proceedings of IEEE Big Data. -2023. DOI:10.1109/bigdata59044.2023.10386254.
4. Li S. et al. Federated and distributed learning applications for electronic health records and structured medical data: a scoping review // Journal of the American Medical Informatics Association. – 2023. – Т. 30. – №. 12. – С. 2041-2049. DOI:10.1093/jamia/ocad170

5. Aghazadeh Ardebili A., Hasidi O., Bendaouia A., Khalil A., Khalil S., Luceri D., Ficarella A. Enhancing resilience in complex energy systems through real-time anomaly detection // Energy Informatics. - 2024. - Vol. 7(1). -P. 96. DOI:10.1186/s42162-024-00401-8.
6. Muratova G., Jumagaliyeva A., Rystygulova V., Abdykerimova E., Turkmenbayev A., Serimbetov B., Yersultanova Z., Omarkulova G. Development of deep learning framework for complex pattern recognition in big data//Eastern-European Journal of Enterprise Technologies. - 2025. -Vol. 6(9 (138)). - P. 54–66. DOI:10.15587/1729-4061.2025.341468.
7. Hovakimyan G., Bravo J. M. Evolving strategies in machine learning: a systematic review of concept drift detection // Information. – 2024. – T. 15. – №. 12. – С. 786. DOI: https://doi.org/10.3390/info15120786
8. Ntumba P., Georgantas N., Christophides V. Adaptive scheduling of continuous operators for iot edge analytics // Future Generation Computer Systems. – 2024. – T. 158. – С. 277-293. DOI: https://doi.org/10.1016/j.future.2024.04.029
9. Mehmood H. et al. A novel edge architecture and solution for detecting concept drift in smart environments // Future Generation Computer Systems. – 2024. – T. 150. – С. 127-143. DOI: https://doi.org/10.1016/j.future.2023.08.023
10. Sulaiman M. et. al. Online deep learning's role in conquering the challenges of streaming data: a survey // Knowledge and Information Systems. – 2025. – T. 67. – №. 4. – С. 3159-3203. DOI: https://doi.org/10.1007/s10115-025-02351-3
11. Rodrigues M.G. et al. A MLOps architecture for near real-time distributed Stream Learning operation deployment // Journal of Network and Computer Applications. – 2025. – T. 238. – С. 104169. DOI: https://doi.org/10.1016/j.jnca.2025.104169
12. Hu L., Lu Y., Feng Y. Concept drift detection based on deep neural networks and autoencoders // Applied Sciences. – 2025. – T. 15. – №. 6. – С. 3056. DOI: https://doi.org/10.3390/app15063056
13. Mahdi O.A. et al. Federated Learning Under Concept Drift: A Systematic Survey of Foundations, Innovations, and Future Research Directions // Electronics. – 2025. – T. 14. – №. 22. – С. 4480. DOI: https://doi.org/10.3390/electronics14224480
14. Rancea A., Anghel I., Cioara T. Edge computing in healthcare: Innovations, opportunities, and challenges // Future internet. – 2024. – T. 16. – №. 9. – С. 329. DOI: https://doi.org/10.3390/fi16090329
15. Suárez-Cetrulo A.L., Quintana D., Cervantes A. A survey on machine learning for recurring concept drifting data streams // Expert Systems with Applications. – 2023. – T. 213. – С. 118934. DOI: https://doi.org/10.1016/j.eswa.2022.118934

Information about the authors

A. Jumagaliyeva – doctoral student, Astana International University, Astana, Kazakhstan; e-mail: jumagalievaainur.m@gmail.com;

M. Kaldarova – PhD, Astana International University, Astana, Kazakhstan; e-mail: mirakaldarova.zh@gmail.com;

R. Ismailova – PhD, Associate Professor, Kyrgyz-Turkish Manas University, Bishkek, Kyrgyzstan; e-mail: rita.ismailova@manas.edu.kg;

E. Abdykerimova – Candidate of Pedagogical Sciences, Professor, Department of Computer Science, Caspian State University of Technology and Engineering named after Sh. Yessenov; e-mail: abdykerimova_el@mail.ru;

A. Turkmenbayev – Candidate of Pedagogical Sciences, Professor. Department of Fundamental Sciences, Caspian State University of Technology and Engineering named after Sh. Yessenov; e-mail: asset.turkmenbaev@yu.edu.kz.