Dynamic Learning for Streaming Data

Nidaa Flaih Hassan[1], Rehab Flaih Hasan[1] , Shaimaa H. Shaker[1]
[1]College of Computer Science, University of Technology, Baghdad, Iraq
Nidaa.f.hassn @uotechnology.edu.iq
Rehab.F.Hassan@uotechnology.edu.iq
Shaimaa.h.shaker@uotechnology.edu.iq

Abstract
Dynamic learning for streaming data has become a crucial strategy for contemporary learning dynamic learning allows models to adjust in real time to non-stationary situations and update progressively an analysis of dynamic learning approaches for instance online and incremental learning drift detection and adaption techniques ensemble tactics active learning and ongoing learning presented in this study . The practical significance of adaptive stream learning is demonstrated by important application domains such as cybersecurity monitoring financial fraud detection smart city and others .

According to the study, real-time adaptability, ongoing knowledge acquisition, scalability, effective resource use, and resilience to idea drift are some of the main benefits of dynamic learning. Nevertheless, a number of drawbacks and restrictions are also mentioned, including model instability, catastrophic forgetting, noise sensitivity, implementation complexity, a lack of historical context, and assessment difficulties. Drift detection, high data velocity, memory limitations, label scarcity, and the stability-plasticity trade-off are important research concerns. The study comes to the conclusion that dynamic learning is essential for real-time AI systems, and that federated learning, edge intelligence, and stronger adaptable models should be the main areas of future research.

Keywords: Dynamic Learning, Streaming Data, Online Learning, Concept Drift, Adaptive Models, Real-Time Analytics, Incremental Learning, Data Stream Mining

1. Introduction
Dynamic learning for streaming data emerged to overcome the limitations of traditional batch machine learning, which assumes static datasets and offline training, conditions unsuitable for modern environments where data arrives continuously and evolves. Early foundations of streaming analytics introduced incremental and one-pass learning approaches to process unbounded data with limited memory, marking a shift from static to real-time learning paradigms [1] .

As research progressed, scholars identified concept drift, the change in statistical properties of data over time, as a major challenge, leading to the development of adaptive techniques such as sliding windows, drift detection algorithms, and incremental decision trees capable of updating models without full retraining [2][3]. Between 2015 and 2020 dynamic learning advanced through an adaptive weighting methods that improved robustness in non-stationary environments And scalable frameworks for real-time stream mining [4][5] as we see the integration of deep learning and continual learning has enabled models to capture temporal patterns forgetting through memory mechanisms and adaptive updates, supporting applications As we see In cybersecurity monitoring, IoT analytics [6]. Current research trends (2023–2026) focus on federated streaming learning, scalable drift adaptation, and lightweight real-time models capable of operating in distributed and resource-constrained environments, highlighting the evolution of dynamic learning into a core technology for modern real-time intelligent systems [7][8].

2. Background and Fundamentals

Predictive models are needed for many real-world applications in the present information technology era, such as network intrusion detection and Self-organized [9][10], financial risk prediction [11], social media sentiment extraction [12], and more. These applications generate an infinite stream of data at great volume and speed [13]. Because the distribution of data in data streams varies, data streams are dynamic. "Concept drift" refers to the shift in the distribution of data [14]. Skewness results from an uneven number of class instances in a data stream [15].Streaming data refers to information that is continuously generated and delivered in a real-time, incremental fashion from multiple sources, which must be processed on-the-fly without storing the entire dataset for later analysis. It is characterized by its high velocity, unbounded nature, and time. sensitive, requiring systems to handle data in real time with limited memory[16]. Streaming data enables immediate insights by making events as soon as they arrive supporting applications as in financial monitoring, sensor networks, and online transactional analytics [17]. Because streaming data cannot be fully stored before processing, stream processing frameworks employ specialized engines and algorithms designed for continuous ingestion, incremental computation, and scalable analytics across distributed environments [18].

Streaming data can be handled using either static or dynamic learning approaches, depending on its characteristics and the application requirements.

Online static learning is one of the first methods studied in online learning [19]. These methods assume nothing changes in the data or features, but in reality, things often change [20][21].

Table (1) shows Types of Streaming Characteristics, Suitable Learning and examples.

Table (1): Types of Streaming Data and Learning Approaches

Type of Streaming Data	Characteristics	Suitable Learning	Examples	References

Sensor / IoT Streams	Continuous, time-stamped, moderate/high velocity	Dynamic Learning	Temperature sensors, smart devices	[16][17]
Transactional Streams	High-volume, structured, frequent updates	Static or Dynamic (if evolving)	E-commerce transactions, bank payments	[17] [18]
Event-based Streams	Irregular, asynchronous, discrete events	Dynamic Learning	Social media posts, system logs	[16][18]
Time-series Streams	Ordered sequences may change gradually	Static (if stationary) / Dynamic (if evolving)	Stock prices, weather data	[16][17] [18]

3. Dynamic Learning Techniques

In order to handle distributional changes and retain performance in changing contexts, dynamic learning for streaming data depends on techniques that enable models to update constantly as new data enters. Rather than retraining on the entire dataset, one fundamental method is online and incremental learning, in which models change their parameters with each new occurrence or small batch. Models can react swiftly to changes like idea drift, which happens when the underlying data distribution changes over time, thanks to this gradual adaptation. Combining latent representation learning with memory modules helps decrease forgetting and initiate updates when drift is identified in developing streams, according to research on adaptive online incremental learning [22].

Combining explicit drift detection and adaptation methods with incremental learning is another crucial tactic. In order to increase predicted accuracy, online classifiers that have been improved using drift detectors and sliding window approaches, for instance, adapt dynamically as patterns change by combining incremental updates with ensemble techniques. In non-stationary settings, these methods offer strong adaptability without sacrificing accuracy[23]**.**

The dynamic learning toolset also includes active and selective learning strategies, especially in situations when labelled data is expensive or scarce. These techniques improve model performance and lower labelling costs in real time by continuously choosing the most instructive cases for labelling as they enter the stream [24] .

Additionally, polls on online streaming and constant learning combine streaming with continual learning principles to retain prior knowledge while adjusting to new information. For long-running systems that need to adapt to current changes while maintaining previous patterns, this is essential [25].

Last but not least, thorough analyses of machine learning for recurrent idea drift outline other adaptive strategies that assist in preserving model relevance over time and enhance resilience against recurrent or cyclical changes in the stream. These strategies include meta learning, model reuse, and ensemble adaptations [26].

4. Applications of Dynamic Learning

Dynamic learning is a potent tool for contemporary applications where real-time decision-making is essential in my opinion it's because of its capacity to manage streaming input . These applications are summarized as follows.

1. Cybersecurity: Intrusion & Anomaly Detection

Cybersecurity uses dynamic learning techniques, especially online and adaptable machine learning models, to quickly identify and address risks when continually examining network traffic these models enhance (IDS) by adjusting to novel attack [22][23][27]. Applications include virus detection, real-time financial network monitoring [23][28][29].

2. Healthcare Monitoring: Wearable Sensors & Patient Tracking

Dynamic learning in healthcare makes it possible to monitor patients in real time using wearables (BANs), and Internet of Things Adoptive algorithms streaming vital sign data to identify abnormalities and notify medical workers promptly, improving patient outcomes [24],[26]. Such techniques work especially well in settings where physiological data, such as heart rate monitoring are changing over time [24].

3. Financial Fraud Detection: Transaction Monitoring

Financial organizations use dynamic learning to analyze transaction in order to detect fraud Compared to static rule-based systems, adaptive algorithms are better able to identify suspicious patterns [26] [30].This involves monitoring credit card and banking activities and increasing detection accuracy [26].

4. Smart Cities: Traffic, Energy & Environmental Systems

Smart city applications rely on dynamic learning to interpret real-time sensor data for environmental monitoring, and traffic optimization. Cities can know circumstances like traffic jams, erratic energy use, or environmental dangers thanks to adaptive algorithms [30][31]. For instance, constant monitoring of electricity grids enhances sustainability and productivity, while spilling traffic data helps optimize signal timings [30].

5. Continuous Knowledge Acquisition cities

This enables systems to react quickly to evolving circumstances, such as novel cyberattacks Models retain current information without extensive retraining. Used in banking systems where fraud tendencies are always changing or in healthcare monitoring, where In smart city applications, where sensors continuously provide traffic

1. **Scalability for Streaming Data**
   Large, fast-moving data streams can be effectively handled using dynamic learning. energy, and environmental data
2. **Efficient Resource Utilization**
   By processing data in tiny batches incremental learning lowers the amount of memory and processing power needed. This benefit is particularly or wearable health monitors, IoT devices, and embedded systems.
3. **Robustness to Concept Drift**
   Adaptive models can recognize and adapt to shifts in the distribution of data.models in cybersecurity continue to stay accurate despite changing fraudulent behaviors

4. **Timely Decision-Making**
   Dynamic learning facilitates quick, well informed decision making by continually updating forecasts .This is crucial for energy grid, traffic control, and medical notifications6. Disadvantages of Dynamic Learning Our investigation revealed the following drawbacks and difficulties with dynamic learning:

1. **Risk of Model Instability**
   - If learning rates or update rules are not properly controlled, continuous updates can cause models to become unstable. It might result in unpredictable forecasts in delicate applications like healthcare monitoring.
2. **Catastrophic Forgetting**
   - Models may lose track of previously learnt knowledge when they adjust to new input.
   - This is particularly troublesome when depend on recurrent patterns (such as seasonal traffic patterns schemes)
3. **Sensitivity to Noise and Outliers**

   - Noisy or inaccurate data might have an immediate impact on the model's performance due to incremental changes.
   - accuracy can be lowered, for example, by aberrant network packets in cybersecurity and sensor issues in smart cities.

4. **Complexity of Implementation**
   - Complex infrastructure is needed for real-time data processing, monitoring,and model.

- This raises the expense of development and upkeep, especially for large-scale applications.

5. **Challenges in Handling Concept Drift**
   - Although dynamic learning can adjust to shifting data distributions, it is still challenging to identify and address many forms of idea drift, such as sudden, gradual, and recurrent.
   - Inaccurate or delayed adaptation might result from improper handling.
6. **Evaluation and Reproducibility Issues**
   - Because performance changes over time, continuous adaptation makes model evaluation more difficult.
   - It becomes difficult to consistently replicate outcomes or benchmarking models, particularly in regulated industries like banking or healthcare.

6. Challenges in Dynamic Learning for Streaming Data

Models can continuously learn from fresh data thanks to dynamic learning for streaming data, but there are a number of difficulties. Real-time processing is challenging due to the speed, volume, and constant change of streaming data. Models have to deal with shifting patterns, memory constraints, noisy data, and occasionally absent labels. Additionally, students should acquire new knowledge without losing sight of what they have already acquired As a result, learning dynamically is more difficult than learning statistics in real-time industries like cybersecurity, healthcare,systems must be quick, Our research on how dynamic learning functions with actual streaming data serves as the base for these problems.

1. Concept Drift: As data distributions change over time, the model's predictions may become erroneous.
2. High Data Velocity: Since streaming data frequently arrives at high rates, models must process and update instantly.
3. Big Data Volume: To prevent system overload, the constant and limitless nature of data streams requires effective memory.
4. Memory Constraints: Learning capacity may be limited because models must learn progressively without keeping all data.
5. Label Scarcity: Supervised learning are difficult in streaming contexts where labelled data may be delayed.
6. (stability) to avoid overlooking important patterns. Noise and Outliers: If handled improperly, mistakes, missing values, or abnormalities found in streaming data can deteriorate model performance.
7. Stability vs. Plasticity Trade-off: To prevent forgetting significant patterns, models must strike a balance between learning (plasticity) and holding (stability)
8. Latency Requirements: most applications need to make predictions very

instantaneously, which can be difficult if the model is complicated or the input stream is moving quickly.

9. Heterogeneous Data Sources: The model must effectively integrate and adapt to data that comes from a variety of sources with different formats Evolving Feature Spaces – New features may appear over time in streaming data, requiring the model to detect and incorporate them without retraining.
10. Unbalanced Data Streams: Rare occurrences of some classes or events might skew learning and lower accuracy for minority classes.
11. Energy and Resource Efficiency: Algorithms must optimize energy consumption and resource allocation since continuous learning are computationally expensive
12. Real-Time Drift Detection: One of the main problems with adaptive models determining when and how the data has changed.
13. Scalability: In distributed or cloud situations, models must continue to function the number of data sources and stream size increase.

7. Limitations of Dynamic Learning for Streaming Data

The potential of dynamic learning for streaming data lies in its ability to instantly adjust to new input. It have certain built-in restrictions . That stem from the characteristics of streaming data, including its fast speed, fluctuating patterns Because of this, dynamic learning may find it difficult to identify long-term patterns deal with drastic changes, or retain complete accuracy in all circumstances. A description of these restrictions is as follows:

1. Limited Historical Context

- Dynamic learning models may not preserve whole history datasets and frequently analyze data.
- It restrict insight for sporadic occurrences in traffic control , finance, or health monitoring by making it harder to identify long-term

2. Trends Model Expressiveness Constraints

- It may be challenging to train complex models incrementally, especially deep neural networks.
- This restricts the kinds of learning algorithms that may be used successfully in situations involving real-time streaming.

3. Dependence on Quality of Data Streams

- Streaming models are totally dependent on the data that comes in. The accuracy and generalizability of the model are intrinsically limited if the data stream is biased, sparse, or incomplete.

4. **Limited Interpretability**
   - Compared to conventional batch-trained models, many dynamic learning techniques, particularly ensemble methods or online deep learning, produce models that are more difficult to understand.
   - This is a drawback in fields like healthcare and financial fraud detection, where the capacity to explain things is crucial.

5. **Inability to Always Handle Extreme Concept Drift**

- while dynamic learning is adapting as I see it can't abrupt, drastic shifts in the data distribution,
- Until the model properly adjusts, it momentarily lower forecast accuracy.

6. **Resource Constraints in High-Velocity Streams**

   - The model's capacity to analyze data in real time may be limited by streaming data at extremely high rates that surpass memory, processing,
   - This is especially important for industrial sensor, smart cities, and Internet of Things networks.

7. **Evaluation Limitations**

The data and model are always changing, it is challenging to assess and evaluate models using evaluation techniques.

8. Conclusion and Future Research Directions

Dynamic learning has emerged as a approach for modern data-driven applications, access models to continuously adapt to evolving streaming data without requiring full retraining. It plays a role in direct decision-making across domains such as finance, healthcare, cybersecurity, and smart cities, as it handles concept high-velocity data—challenges that static learning methods struggle to address. study the differences between static and dynamic learning and reviews strategies, including online learning, incremental learning, and adaptive techniques. It also explores mechanisms for drift detection and model adaptation.

The findings indicate that dynamic learning provides , scalable, and cost-efficient solutions that ensure robustness and predictive accuracyin environments There are still crucial directions for future study in dynamic learning, despite advancements. streaming learning presents the possibility of distributed model updates across several devices or locations while maintaining anonymity. In order to lower latency and improve responsiveness, AI at the edge focuses on implementing adaptive models near the data source. Performance in challenging streaming contexts enhanced by hybrid learning that include online, incremental, self- adaptive intelligent systems capable of automatically adjusting their strategies in response to data changes represent a promising direction for achieving autonomous, real-time AI systems. Pursuing these will strengthen the of dynamic learning as a component of modern AI and real-time analytics..

References

[1] A. H. Farhan, O. S. F. Shareef, and R. F. Hasan, "The Effect of False Predictions of Machine Learning on the Security of the Big Data Environment", Iraqi Journal of Science, 2025, Vol. 66, No. 1, pp: 361-374,2025.

[2] J. Gama, I. Žliobaitė, A. Bifet, M. Pechenizkiy, and A. Bouchachia, A. , " A survey on concept drift adaptation. ACM Computing Surveys" , 2013.

[3] S. Mouchaweh, "Learning from Data Streams in Dynamic Environments", SpringerBriefs in Computer Science,2015.

[4] A. Bifet, G. Holmes, B. Pfahringer, and C. Giraud-Carrier,"Machine Learning for Data Streams with Practical Methods". MIT Press.2018

[5] J. Lu, A. Liu, F. Dong, F. Gu, J. Gama and G. Zhang , " Learning under concept drift: A review", IEEE Transactions on Knowledge and Data Engineering , 2018.

[6] A. J.Rabash, andD. Maltoni ," Stream learning under concept and feature drift: A literature survey. Journal of Autonomous Intelligence", 2023.

[7]T. W.Gheed and S. H.Shaker, " Speech Emotion Recognition on MELD and RAVDESS Datasets Using CNN", Information , vol 16, no. 518, 2025,.

[8] X. Chen, S. Wang, and T. Zhang, "Virtual concept drift detection and adaptation in federated streaming learning. International Journal of Data Science and Analytics", 2025.

[9] M. A. Rezvi, S. Moontaha, K. A. Trisha, S. T. Cynthia, and S. Ripon, "Data mining approach to analyzing intrusion detection of wireless sensor network," Indonesian Journal of Electrical Engineering and Computer Science, 2021 .

[10] N. G. Adiel, A. A. Salman, and N. F. Hassan, "Node Failure in Self-Organized Sensor Networks," Open Engineering, vol. 15, no. 1, Art. no. 20240083, 2025.

[11] J. Sun, H. Li, H. Fujita, B. Fu, and W. Ai, "Class-imbalanced dynamic financial distress prediction based on Adaboost-SVM ensemble combined with SMOTE and time weighting," Information Fusion, vol. 54, pp. 128–144, Feb. 2020, doi: 10.1016/j.inffus.2019.07.006

[12] N. Seman and N. AtiqahRazmi, "Machine learning-based technique for big data sentiments extraction," IAES International Journal of Artificial Intelligence (IJ-AI), vol. 9, no. 3, pp. 473–479, Sep. 2020, doi: 10.11591/ijai.v9.i3.pp473-479.

[13] E. Alothali, H. Alashwal, and S. Harous, "Data stream mining techniques: a review," TELKOMNIKA, vol. 17, no. 2, pp. 728–737, 2019, doi: 10.12928/TELKOMNIKA.v17i2.11752.

[14] L. Deshpande and M. N. Rao, "Concept drift identification using classifier ensemble approach," International Journal of Electrical and Computer Engineering (IJECE), vol. 8, no. 1, pp. 19–25, Feb. 2018, doi: 10.11591/ijece.v8i1.pp19-25.

[15] S. Ancy and D. Paulraj, "Handling imbalanced data with concept drift by applying dynamic sampling and ensemble classification model," Computer Communications, vol. 153, pp. 553–560, 2020, doi: 10.1016/j.comcom.2020.01.061.

[16] T. Kolajo, O. Daramola, and A. Adebiyi, "Big data stream analysis: a systematic literature review", Journal of Big Data, vol. 6, Art. 47, 2019. https://journalofbigdata.springeropen.com/articles/10.1186/s40537-019-0210-7

[17] M. A. Alam, A. R. Nabil, A. A. Mintoo, and A. Islam, "Real-Time Analytics in Streaming Big Data: Techniques and Applications", Journal of Science and Engineering Research, vol. 1, no. 01, pp. 104–122, 2024. https://www.researchgate.net/publication/386283651

[18] U. Kekevi and A. Aydın, "Real-Time Big Data Processing and Analytics: Concepts, Technologies, and Domains, Computer Science", vol. 7, no. 2, pp. 111–123, 2022. https://www.researchgate.net/publication/366071082

[19] M. Millan-Giraldo, and J.S. Sanchez,"Online Learning Strategies for Classification of Static Data Streams", in Proceedings of the 8th WSEAS International Conference on Multimedia, 2008.

[20] H. Lian, D. Wu, D. Hou, J. Wu, and Y. He, Online Learning from Evolving Feature Spaces with Deep Variational Models. IEEE Trans. Know. Data Eng. 2023, 36, 4144–4162.

[21] Y. He,B. Wu, D. Wu, E. Beyazit, S. Chen, S., and X. Wu, "Toward Mining Capricious Data Streams: A Generative Approach". IEEE Trans. Neural Netw. Learn. Syst. 2021, 32, 1228–1240.

[22] S. Zhang, J. Liu, and X. Zuo, "Adaptive online incremental learning for evolving data streams," Applied Soft Computing, vol. 105, art. 107255, Jul. 2021.

[23]D. Mulimani ,P. Patil , S. Totad , and R. Benni, "Online detection and adaptation of concept drift in streaming data classification," in Proc. of the International Conference on Advances in Data Stream Mining and Processing, Procedia Computer Science, vol. 235, pp. 2803–2811, 2024.

[24] D. Cacciarelli and M. Kulahci, "Active learning for data streams: a survey," Machine Learning, vol. 113, pp. 185–239, 2024.

[25] N. Gunasekara, B. Pfahringer, H. M. Gomes, and A. Bifet, "Survey on online streaming continual learning," in Proc. 32nd International Joint Conference on Artificial Intelligence (IJCAI 2023), Macao, China, pp. 6628–6637, 2023.

[26] A. L. Suárez-Cetrulo, D. Quintana, and A. Cervantes, "A survey on machine learning for recurring concept drifting data streams," Expert Systems with Applications, vol. 213, art. 118934, 2022

[27] S. F.Abbas, S. H. Shaker ,and F. A. Abdullatif," Masked Face Identification and Tracking Using Deep Learning", Journal of Computer Science , 19 (12): 1423.1437 , 2023
DOI: 10.3844/jcssp.2023.1423.1437.

[28] A. C. Ahmed, M. A. Khan, and R. Ali, "Real-time anomaly detection in cybersecurity using online learning," Journal of Information Security, vol. 14, pp. 45–62, 2023.

[29] A. S. Mohamed, N. F. Hassan, and A. S. Jamil, "Real-Time Hand Gesture Recognition: A Comprehensive Review of Techniques, Applications, and Challenges," Cybernetics and Information Technologies, vol. 24, no. 3, pp. 163–181, 2024.

[30] A. Zanella et al., "Internet of Things for Smart Cities", IEEE Internet of Things Journal, vol. 1, no. 1, pp. 22–32, Feb. 2014.

[31] S. E. Shafiee, M. R. Ahmad, and M. Mohammadi, "Smart city applications and challenges: A survey," Sustainable Cities and Society, vol. 63, art. 102471, 2021.