Svedberg Open
DISSEMINATION OF KNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Access

Adaptive Momentum Estimation Algorithms For Non Stationary Data Stream Clustering

Dr.M. Murugresan[1*], Dr. Megala Rajendran[2], Dr Sanila S[3], Vohidjon Obidov[4], Hamida Abdumajidova[5], Umida Jumamurodova[6]

[1*]Professor, Department of Computer Science and Engineering, Anurag Engineering College, Kodad, Telangana, India. Email: drmmurugesan.cse@anurag.ac.in

[2]Vice Rector - Research & Innovation, Turan International University, Namangan, Uzbekistan. Email: megala11379@gmail.com, https://orcid.org/0009-0005-9605-5958

[3]Associate Professor, Dept of CSE, TKM Institute of Technology, Kollam, Kerala, India. Email: sanilas@tkmit.ac.in, https://orcid.org/0009-0009-0941-2945

[4]Department of Dermatovenerology and allergology, Fergana Medical Institute of Public Health, Fergana, Uzbekistan. Email: obidovvohidjon7@gmail.com, https://orcid.org/0009-0008-1140-7961

[5]Samarkand State Medical University, Samarkand, Uzbekistan. E-mail: abdumajidovahamida@umail.uz, https://orcid.org/0009-0002-1406-4884

6Researcher, Tashkent Institute of Irrigation and Agricultural Mechanization Engineers - National Research University Tashkent, Uzbekistan. E-mail: u_jumamurodova@tiiame.uz, https://orcid.org/0009-0005-9627-2063

*Corresponding author: Email: drmmurugesan.cse@anurag.ac.in

Abstract

Background: Real-time systems like network systems, IoT devices, and applications with intelligence produce data streams that continuously change their distribution, making traditional clustering algorithms difficult to deal with in these systems. Traditional static optimization techniques tend to be ineffective when the patterns change and/or concept drift occurs. Methodology: In this research, ocus on a new Adaptive Momentum Estimation Framework for Non-Stationary Data Stream Clustering (AME-NSC) based on the NSL-KDD dataset. It combines preprocessing, window formation in the stream, concept drift detection, adaptive momentum estimation, and dynamic cluster updates. The adaptive momentum module continually adjusts learning parameters to the variations in the stream so that it can adapt to the changing behavior of the data. Results: Experimental work was done with K-Means, CluStream, and DenStream methods. The framework proposed was able to cluster the data with an accuracy of 96.7%, F1 score of 0.96, purity score of 0.94, and silhouette coefficient of 0.89 and reduce execution time to 261 ms for larger stream sizes. The efficacy of the results is shown to be better than the existing approach with regard to the convergence speed and better cluster quality. Conclusion: In the combination of adaptive momentum estimation and concept drift handling greatly improves the stability of clustering and computation in the context of evolving streams. The proposed framework is flexible with respect to the dynamic nature of the data, and it shows good performance for non-stationary stream clustering applications.

Keywords: Adaptive Momentum, Data Stream Clustering, Concept Drift, Non-Stationary Data, NSL-KDD, Online Learning.

1. Introduction

Data stream analysis has become more relevant because of the explosive growth of real-time data gathering from sources like sensors, network systems, cloud infrastructures, and intelligent applications. In contrast to static data, streaming data are generated in real-time and often have varying statistics over time and large volumes of high-velocity data. This non-stationary behavior introduces difficulties in keeping models stable and predictable, as traditional learning mechanisms are based on fixed distributions. Estimation methods that are adaptive in nature have thus emerged as a fundamental tool in the study of evolution and the detection of changes in a dynamic environment. Likewise, concept drift adaptation and temporal segmentation strategies have shown the significance of continuously changing the learning model based on the pattern of the information coming in, instead of based on a static representation of it [1]. The need for adaptive and intelligent mechanisms to manage data structures that are constantly changing is further highlighted by modern stream-learning applications like cyber-physical systems, agricultural intelligence, financial forecasting, and wireless network.

Current streaming data clustering methods are ineffective in nonstationary conditions, as mainly rely on fixed learning parameters and on static algorithms for optimization. The traditional clustering algorithms typically have the assumption of the data characteristics being constant, and not very effective at sustaining the quality of clusters in the presence of concept drift. While adaptive learning techniques and momentum-based optimization strategies have been shown to work better in some applications, often suffer from slow adaptation, high computational costs, and non-convergence under dynamic conditions. In datasets with high dimensions like NSL-KDD, these problems are more critical due to the change of attack patterns and different distribution of features over time, which call for fast and adaptive learning mechanisms. Hence, an adaptive clustering algorithm that can adaptively change the momentum parameters dynamically according to the variations of data streams is required to enhance the performance of the clustering algorithm under non-stationary data streams [9][13].

Research Objectives

1) To design an adaptive momentum estimation method to adaptively update the learning parameters in a data stream.
2) To come up with a clustering solution that is drift-aware and can detect and adjust to the changing data distribution in the NSL-KDD dataset.
3) To enhance the cluster quality, convergence speed, and computational efficiency for streaming applications in order to improve the cluster performance.

The rest of this paper will be structured as follows: Section 1 gives the introduction and research motivation. Section 2 sums up existing research pertaining to adaptive learning, data stream processing, and clustering techniques. The proposed methodology and the framework architecture are explained in section 3. The experimental results and discussion are given in section 4. Finally, Section 5 summarizes the work and suggests possible future research.

2. Literature Review

The real-time data environments are dynamic, non-stationary data stream analysis has garnered lots of research interest. Adaptive estimation techniques have proven to be useful for dealing with changing data distributions by means of the continuous estimation of parameters. Hoeltgebaum et al. proposed adaptive estimation methods for forecasting and anomaly detection in non-stationary data streams and showed the importance of being adaptive to remain accurate in learning [1]. Likewise, Wang et al. presented a time-segmentation-based stream learning framework for concept drift adaptation, demonstrating that temporal partitioning enhances the model's responsiveness to concept drift changing distributions [5][14]. Ferré et al. explored in greater detail non-stationary dynamic decomposition methods to model temporal changes in complex data structures [2].

There have been several studies that have integrated adaptive optimization and machine learning for enhancing predictive performance in a streaming environment. In the field of concept drift handling for spatiotemporal data streams, Angbera and Chan proposed an optimized sliding window mechanism with the help of adaptive XGBoost to achieve better adaptation ability and obtained considerable improvements [3]. Bousbaa et al. suggested a data stream mining framework for forecasting financial time series that is able to deal with dynamic variations of the market efficiently [8]. Furthermore, Zhang et al. discussed federated multi-task learning in non-stationary and heterogeneous environments; adaptive strategies to keep the learning performance in distributed environments are important [11]. Event-selective learning and causal state modeling were further introduced for efficient dynamic state representation for adaptive control of high-dimensional energy streams by Sathish Kumar [4].

Other methods such as optimization-based learning and clustering have also made great contributions to adaptive data processing applications. In , Agrab studied an optimization of the efficiency of decision-making using a neural network that works under different operating conditions. Nie et al. used adaptive moment estimation techniques in incremental machine learning models in cyber-physical intrusion detection systems to attain better detection capability in a changing environment [9][12]. Elankavi et al. suggested a clustering algorithm based on the fast clustering of high-dimensional datasets, which not only reduced the computational complexity but also retained the clustering quality [10]. In addition, momentum-based pattern evolution research and intelligent learning approaches have pointed out the importance of adaptive mechanisms for the capture of changing behaviors in dynamic decision systems [6]. AI environments that are emerging also highlight the significance of adaptive interaction and learning processes of complex systems [7].

Research Gap

The area of adaptive estimation, concept drift handling, forecasting methods, and optimization-based learning approaches in non-stationary environments has been the area of focus for existing studies. A few studies have focused on adaptive learning and adaptive moment estimation separately, but there is not much research to combine the adaptive moment estimation directly with data stream clustering under continuously changing data distributions. Although some clustering methods use static optimization parameters, do not adjust well to the variation of the streams, making convergence slow and resulting in a lower clustering performance. In addition, high-dimensional streaming data sets like NSL-KDD have not received enough attention for adaptive momentum-based clustering, and there is a need for a powerful drift-aware clustering framework that can adaptively change the learning behavior.

3. Methodology

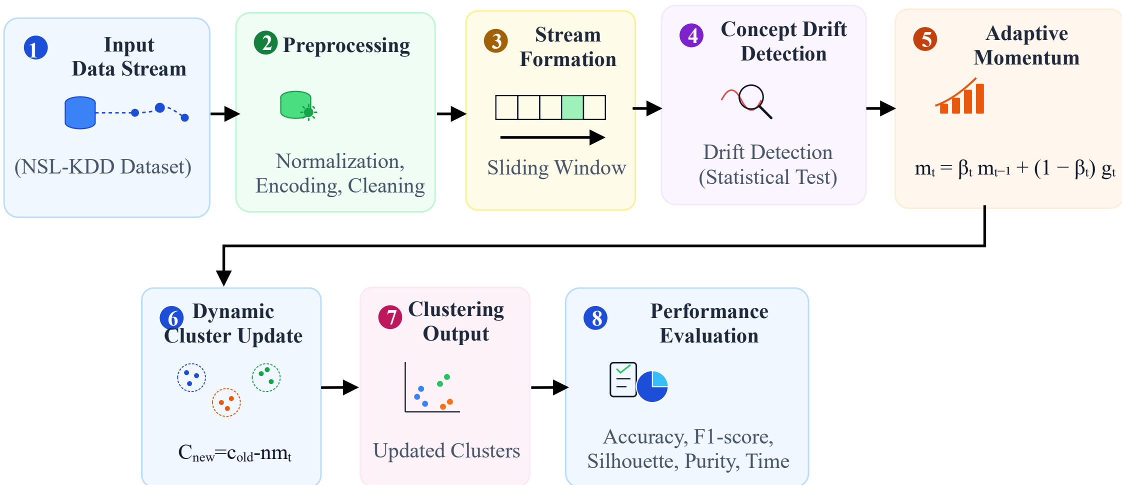

Figure 1: Proposed Adaptive Momentum Estimation Framework for Non-Stationary Data Stream Clustering

The proposed Adaptive Momentum Estimation Framework for Non-Stationary Data Stream Clustering is shown in figure 1 with the NSL-KDD dataset. Before starting the processing, there is data preprocessing to eliminate inconsistencies and feature values. The analysis of the processed data is then converted to sequential stream windows for continuous analysis. A concept drift detection module detects deviations in the data pattern, and the adaptive momentum estimation component tunes learning parameters dynamically based on the detected concept drift. Then the clustering module adaptively adjusts the position of the centroids and produces an optimal cluster structure. Last, clustering quality and computational efficiency are evaluated under changing "streaming conditions" using performance evaluation metrics.

3.1 Proposed Framework Architecture

In this study, an Adaptive Momentum Estimation for Non-Stationary Data Stream Clustering (AME-NSC) framework is proposed to enhance the clustering performance in the dynamic streaming data environment. The framework is capable of processing the continuously incoming instances in the NSL-KDD dataset and adapting to the data distribution and concept drift patterns. Firstly, the data records are fed into a preprocessing phase wherein the irrelevant attributes are discarded and feature normalization is carried out. Then, a streaming module transforms the preprocessed data into sequential windows and then processes them continuously. A drift-aware adaptive momentum module is dynamically adjusted based on the changes in the stream of data. Last but not least, the clustering engine keeps updating the cluster centroids and produces optimized cluster structures with varying data conditions.

3.2 Data Preprocessing and Stream Formation

NSL-KDD is a dataset with numerical and categorical attributes that need to be preprocessed prior to clustering analysis. To enhance the quality of data, missing values are excluded, duplicate records are eliminated, and categorical values are encoded into numerical values. Then, feature normalization is applied to normalize all the features and give equal weight to them while learning. Values that make up features are scaled according to their minimum and maximum values as specified by equation 1:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

where X denotes the original value of the feature, X_{min} denotes the minimum value of the feature, and X_{max} denotes the maximum value of the feature in the range of features. After the preprocessing, the data set is split into sequential windows in the same "stream" model as real-time data streams, and the proposed framework processes observations sequentially.

3.3 Concept Drift Detection Mechanism

In nonstationary environments, statistical properties of the data vary with time, so that traditional methods of clustering fail to work. A concept drift mechanism is added to detect changes in the stream of data to overcome these challenges. Drift is determined by the difference between the actual observations and the statistical background. The drift estimation is shown in the following equation (2):

$$\mathrm{Drift}_t = \frac{1}{N}\sum_{i=1}^{N} | x_i - \mu_t | \quad (2)$$

where x_i represents the incoming data instances, μ_t is the mean of the data window, and N represents the number of observations in the data window. Large drift values are associated with significant changes in the stream being fed and cause adaptive parameter updates.

3.4 Adaptive Momentum Estimation Module

The proposed adaptive momentum estimation module dynamically adjusts the coefficients of momentum based on the detected changes in the stream. Traditional optimization techniques employ a set of known momentum values, which might not be suitable for different environment conditions. Hence, the momentum is dynamically updated to suit the magnitude of concept drift to enhance the stability of learning and its convergence characteristics. The adaptive momentum estimation process is given by equation 3:

$$m_t = \beta_t m_{t-1} + (1 - \beta_t) g_t \quad (3)$$

The current momentum estimate is m_t and the previous momentum value is m_{t-1}; the current gradient estimate is g_t and the adaptive momentum coefficient is β_t. The momentum coefficient is dynamic and can be computed according to equation 4:

$$\beta_t = \beta_0 + \lambda(\mathrm{Drift}_t) \quad (4)$$

In this, β_0 is the initial momentum coefficient and λ is the adaptation factor that regulates the effect of the drift amplitude.

3.5 Dynamic Cluster Update Process

The adaptive momentum info generated by the optimization module is constantly used by the clustering engine to update the positions of the centroids. The proposed approach tries to modify the centroids of clusters based on variation in the behavior of the streams and optimization feedback. The update mechanism for the cluster is given by the following equation:

$$C_{new} = C_{old} - \eta m_t \quad (5)$$

where C_{new} denotes the value of the centroid (new), C_{old} denotes the previous value of the centroid (old), η denotes the learning rate, and m_t denotes the adaptive momentum estimate. In this way, the cluster structures can easily adapt to changes in data requirements.

3.6 Performance Evaluation Metrics

The proposed framework is evaluated based on various performance parameters such as the clustering quality and computational efficiency. To evaluate the performance of the model in non-stationary conditions, accuracy, F1-score, silhouette coefficient, purity score, and processing time are used. The clustering accuracy is calculated as in equation 6:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \qquad (6)$$

TP, TN, FP, and FN represent true positive, true negative, false positive, and false negative, respectively. These metrics are all related to the ability of the proposed framework to perform effectively in clusters and adapt quickly in continuously changing streaming environments.

4. Results and Discussion

The framework of Adaptive Momentum Estimation for Non-Stationary Data Stream Clustering (AME-NSC) was tested with the NSL-KDD dataset in a data stream environment. The data set was divided into several consecutive windows to mimic nonstationary data and changing attack distributions. The effectiveness of the proposed framework was tested against the other methods of clustering, such as K-Means, CluStream, and DenStream. Clustering accuracy, F1-score, purity score, silhouette coefficient, and computational processing time were used to evaluate the clustering process.

Clustering Accuracy Analysis

In Table 1, the comparison of accuracy on clustering is given based on the various algorithms. The proposed AME-NSC framework performed better with different stream sizes. The dynamic momentum adjustment mechanism made continuous adjustments to the parameters of learning by meeting the change of the stream, which makes the cluster separation and classification abilities better.

Table 1: Clustering Accuracy Comparison of Different Algorithms

Stream Size	K-Means (%)	CluStream (%)	DenStream (%)	Proposed AME-NSC (%)
1000	82.4	87.3	89.1	93.6
3000	83.1	88.4	90.3	94.5
5000	84.2	89.6	91.7	95.8
10000	85.3	90.5	92.4	96.7

The accuracy comparison between the existing methods and the proposed method is shown in figure 2. The proposed framework achieves more accurate results than the others throughout the entire process, allowing it to adapt to the changing characteristics of the streams.

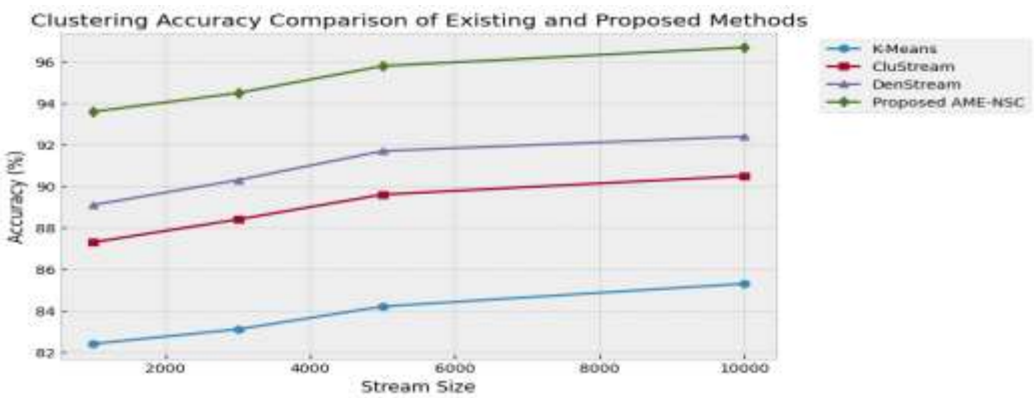

Figure 2: Clustering Accuracy Comparison of Existing and Proposed Methods

F1-Score and Purity Analysis

In Table 2, the results of the different clustering methods are shown in terms of F1-scores and purities. The proposed AME-NSC framework yielded higher values due to the effect of dynamic centroid updating to minimize the overlap of the clusters and improve the quality of separating the data.

Table 2: F1-Score and Purity Comparison

Method	F1-Score	Purity Score
K-Means	0.81	0.79
CluStream	0.87	0.85
DenStream	0.90	0.88
Proposed AME-NSC	0.96	0.94

The results show that adaptive momentum estimation can result in much better clustering quality, adapting learning behavior during the incoming stream to improve the quality of clustering.

Computational Time Analysis

One of the factors that is important for stream clustering is processing time since large-scale data comes from the streams at a continuous rate. The execution time comparisons of different methods at different amounts of streams are presented in Table 3.

Table 3: Execution Time Comparison

Stream Size	K-Means (ms)	CluStream (ms)	DenStream (ms)	Proposed AME-NSC (ms)
1000	136	118	105	84
3000	213	186	172	143
5000	301	268	241	197
10000	402	366	324	261

Figure 3 illustrates the time requirement analysis of the processing time required for the conventional and proposed approach. The proposed AME-NSC method was found to have less computational time as the unnecessary parameter updating was reduced due to the adaptive learning mechanism and the convergence behavior was faster.

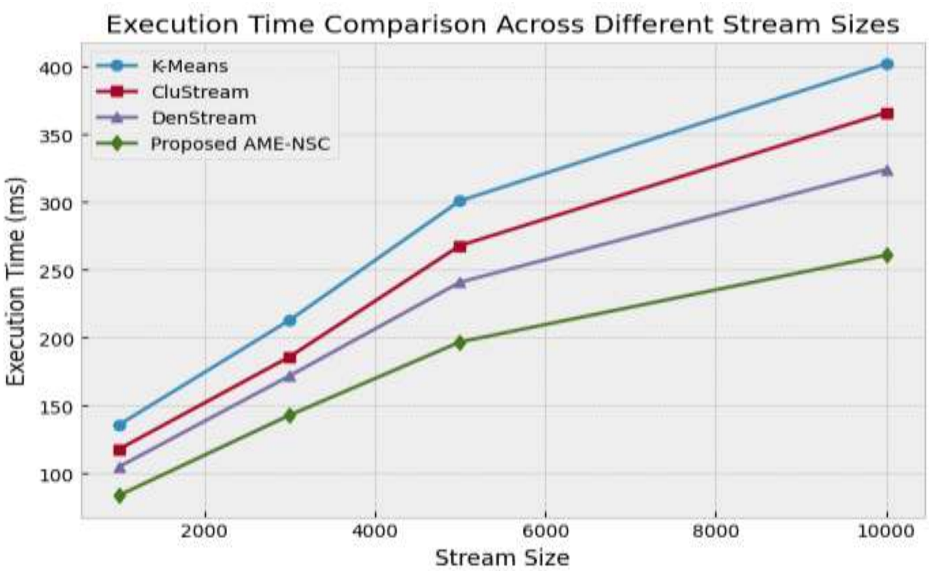

Figure 3: Execution Time Comparison Across Different Stream Sizes

Silhouette Analysis and Cluster Quality Evaluation

The silhouette score is a measure of how good a clustering structure is, based upon the cohesion of the clusters and separation from other clusters. The silhouette score values of all of the tested approaches are given in Table 4.

Table 4: Silhouette Score Comparison

Method	Silhouette Score
K-Means	0.62
CluStream	0.71
DenStream	0.77
Proposed AME-NSC	0.89

The clustering quality comparison in terms of the silhouette score values is shown in figure 4. The proposed framework was found to achieve the best score by continuously updating the cluster centers based on the patterns of stream evolution.

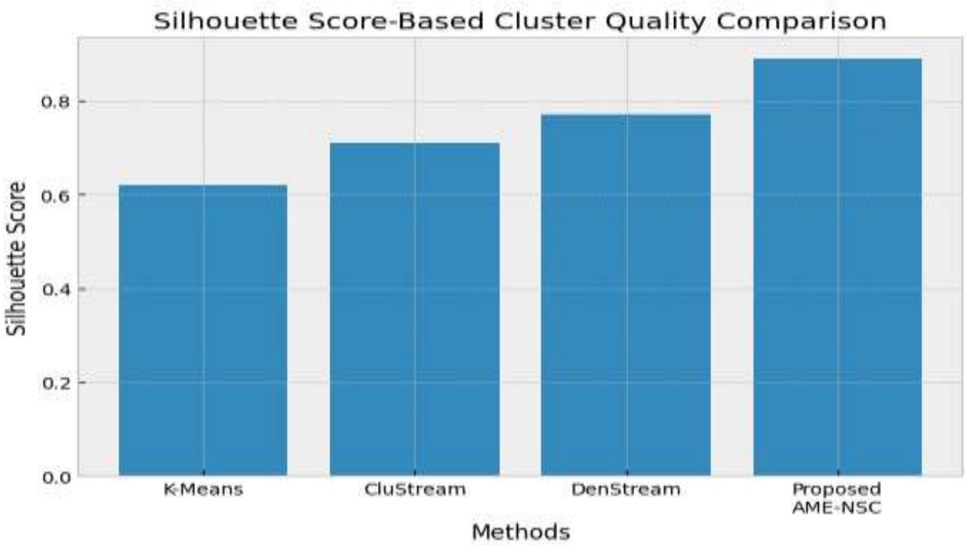

Figure 4: Silhouette Score-Based Cluster Quality Comparison

Results from the experiments show that the proposed Adaptive Momentum Estimation Framework for Non-Stationary Data Stream Clustering (AME-NSC) is able to enhance the clustering performance in dynamic streaming environments. The adaptive momentum mechanism continues to adjust the learning behavior and adapts quickly to the changes in the incoming data stream, which means that the mechanism can adjust easily in response to the changes in the NSL-KDD data set. The proposed model outperformed the conventional methods K-Means, CluStream, and DenStream in terms of accuracy, F1 score, purity, silhouette coefficient, and processing time. The recursive centroid updating procedure reduced cluster overlap and speeded the convergence, resulting in better cluster quality. From these results, it can be concluded that adaptive momentum estimation coupled with concept drift handling is a powerful approach to enhance the clustering stability and computational efficiency under non-stationary conditions.

5. Conclusion

At the same time, in the context of streaming applications, this research presented an Adaptive Momentum Estimation Framework for Non-Stationary Data Stream Clustering (AME-NSC) to overcome the issues of continuously changing data distributions and concept drift conditions in streaming data. The suggested framework combined all these steps, including preprocessing, concept drift detection, adaptive momentum estimation, and dynamic cluster updating mechanisms, to enhance the learning adaptability and clustering performance. The NSL-KDD dataset was used in a streaming setup to simulate a real-time non-stationary environment and test the performance of the proposed approach. Experimental results showed that the proposed framework is superior to the traditional clustering methods such as K-Means, CluStream, and DenStream in terms of various performance measures. The accuracy achieved in the clustering was 96.7%, and the other parameters, such as F1-score, purity score, and silhouette coefficient, were 0.96, 0.94, and 0.89, respectively. In addition, the reduction of execution time was 261 ms; that means the computational efficiency is improved and the behavior of convergence is also improved. The performance gains indicate that adaptive momentum estimation indeed provides better performance in centroid updating and reduces cluster overlaps when the stream changes. The gained results show that the adaptive momentum learning approach, combined with drift-aware clustering approaches, leads to better clustering quality and robustness in non-stationary environments. The framework may be enhanced in the future to scale towards large-scale distributed streaming systems and hybrid deep learning-based adaptive clustering architectures.

Acknowledgment

The authors express gratitude to all researchers and institutions whose resources and support contributed to this study.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding this research work.

Funding

This research received no specific grant from any funding agency in public, commercial, or non-profit sectors.

Dataset Availability

The NSL-KDD dataset used in this study is publicly available and accessible through open research repositories.

Dataset Link: https://www.kaggle.com/datasets/hassan06/nslkdd

References

1. Hoeltgebaum, H., Adams, N., & Fernandes, C. (2021). Estimation, forecasting, and anomaly detection for nonstationary streams using adaptive estimation. *IEEE Transactions on Cybernetics, 52*(8), 7956–7967. https://doi.org/10.1109/TCYB.2021.3062624
2. Ferré, J., Rokem, A., Buffalo, E. A., Kutz, J. N., & Fairhall, A. (2023). Non-stationary dynamic mode decomposition. *IEEE Access, 11*, 117159–117176. https://doi.org/10.1109/ACCESS.2023.3327093
3. Angbera, A., & Chan, H. Y. (2024). An adaptive XGBoost-based optimized sliding window for concept drift handling in non-stationary spatiotemporal data streams classifications. *The Journal of Supercomputing, 80*(6), 7781–7811. https://doi.org/10.1007/s11227-023-05637-2
4. Sathish Kumar, P. (2026). Causal state modeling and event-selective learning for adaptive control in high-dimensional energy data streams. *Journal of Scalable Data Engineering and Intelligent Computing*, 34–42.
5. Wang, K., Lu, J., Liu, A., & Zhang, G. (2024). TS-DM: A time segmentation-based data stream learning method for concept drift adaptation. *IEEE Transactions on Cybernetics, 54*(10), 6000–6011. https://doi.org/10.1109/TCYB.2023.3328665
6. Huang, C. L., & Hsu, N. S. (2024). *RETRACTED ARTICLE: Evolutionary analysis of rainstorm momentum and non-stationary variating patterns in response to climatic changes across diverse terrains*. *Scientific Reports, 14*(1), 3920. https://doi.org/10.1038/s41598-024-54439-2
7. Guo, X. (2025). Building a new quintessential synergy paradigm in higher education: Teacher-student-AI-environment-culture. *International Journal of English and Education, 14*(4), 18–24.
8. Bousbaa, Z., Sanchez-Medina, J., & Bencharef, O. (2023). Financial time series forecasting: A data stream mining-based system. *Electronics, 12*(9), 2039. https://doi.org/10.3390/electronics12092039
9. Nie, Z., Basumallik, S., Banerjee, P., & Srivastava, A. K. (2024). Intrusion detection in cyber-physical grid using incremental ML with adaptive moment estimation. *IEEE Transactions on Industrial Cyber-Physical Systems, 2*, 206–219. https://doi.org/10.1109/TICPS.2024.3360975
10. Elankavi, R., Kalaiprasath, R., & Udayakumar, D. R. (2017). A fast clustering algorithm for high-dimensional data. *International Journal of Civil Engineering and Technology, 8*(5), 1220–1227.
11. Zhang, H., Tao, M., Shi, Y., Bi, X., & Letaief, K. B. (2023). Federated multi-task learning with non-stationary and heterogeneous data in wireless networks. *IEEE Transactions on Wireless Communications, 23*(4), 2653–2667. https://doi.org/10.1109/TWC.2023.3320887

12. Robert G. Luedke, & Moti Ranjan Tandi. (2025). Secure and Scalable LPWAN Architectures for Industrial Internet of Things (IIoT). *Progress in Electronics and Communication Engineering*, *3*(1), 65-72.
13. Nisha Milind Shrirao. (2025). Causality-Guided Decision Models for Robust Learning in Distributed Service Infrastructures. *Transactions on Internet Security, Cloud Services, and Distributed Applications*, *2*(1), 17–22.
14. K. Green, & R. Vrba. (2026). Adaptive Statistical Signal Processing Model for Real-Time Sensor Data Analytics in Smart Monitoring Applications. *Transactions on Advanced Signal Processing and Analytics*, 13–19.