ASEC 2025 Conference
The 6th International Electronic Conference on Applied Sciences
09–11 December 2025 | Online
MDPI

Adaptive Multimodal LSTM with Online Learning for Evolving IoT Data Streams

Osaretin Edith Okoro, Nurudeen Mahmud Ibrahim, Prema Kirubakaran and Suleiman Aliyu Muhammad
Nile University of Nigeria

INTRODUCTION

The rapid growth of the Internet of Things (IoT) is generating vast, heterogeneous, and continually evolving data from sensors, video streams, and network logs. A major analytical challenge is concept drift, where shifting data distributions reduce the accuracy of AI models trained on past conditions. While Long Short-Term Memory (LSTM) networks are widely used for temporal modelling, their limited adaptability makes them insufficient for dynamic multimodal IoT streams. Existing methods address drift detection or multimodal fusion separately. This study proposes the Adaptive Multimodal LSTM (AM-LSTM), a unified framework that integrates drift-aware adaptation and dynamic fusion, improving robustness and overall performance.

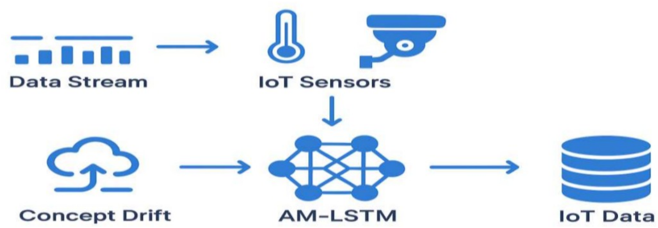

Figure 1: Adaptive Multimodal LSTM (AM-LSTM)

AIM

The aim of this study is to develop and evaluate an Adaptive Multimodal LSTM (AM-LSTM) framework designed to improve real-time prediction in dynamic IoT environments. The study addresses key challenges such as concept drift, heterogeneous data streams, and missing or incomplete modalities. By integrating drift-aware online adaptation, modality-specific temporal modelling, and dynamic multimodal fusion, the AM-LSTM enhances the robustness and accuracy of IoT data analysis. The proposed framework seeks to overcome the limitations of traditional static LSTM models and deliver superior performance across evolving IoT contexts.

METHODOLOGY

The proposed study suggests the Adaptive Multimodal LSTM (AM-LSTM) framework to combine drift-aware online adaptation, attention-based fusion, and modality-specific t emporal modelling to continuously learn from changing IoT streams

Modality Specific Encoding

$$h_t^{(m)}, c_t^{(m)} = LSTM\left(x_t^{(m)}, h_{t-1}^{(m)}, c_{t-1}^{(m)}, \theta^{(m)}\right) \qquad (1)$$

where $h_t^{(m)}$ and $c_t^{(m)}$ are the hidden and cell states, and $\Theta^{(m)}$represent learnable parameters. This allows modality-specific temporal dependencies to be captured independently.

Dynamic Multimodal Fusion

Previous studies [18], [19] on multimodal LSTM integrate the encoded representations attention-based fusion operator:

$$z_t = \sum_{m=1}^{M} \alpha_t^{(m)} h_t^{(m)}, \alpha_t^{(m)} = \frac{exp\left(W_f h_t^{(m)}\right)}{\sum_{j=1}^{M} exp\left(W_f h_t^{(m)}\right)}, \qquad (2)$$

Where $\alpha_t^{(m)}$ represents adaptive modality weights, enabling the framework to emphasize informative streams and down-weight noisy or missing modalities.

Drift Adaptation and Online Learning

Our framework makes use of a sliding window to track prediction error for the management of the changing data distributions. Adaptation is implemented by the framework when error surpass a set threshold using recent samples to selectively retrain LSTMs specific to the impacted modality. In addition, the learning rate is dynamically adjusted during drift phases. To prevent error propagation, models with larger shift should have their states set.

Computational Efficiency

From initial tests, AM-LSTM always maintains updated latency below 50ms by using lightweight update this routine similar to packet LSTM is appropriate given the resource limitations of IoT devices. This framework can be deployed in real-time IoT environment because it strikes a balance between accuracy, adaptability and efficiency.

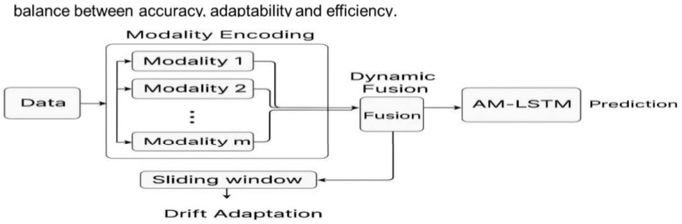

Figure 2: Adaptive Multimodal LSTM (AM-LSTM) framework

RESULTS

Model	Accuracy (%)	F1 – Score	Drift Adapt. Latency Sample	Update Latency (ms)
Static LSTM	72.3	0.69	>2000	35
Packet LSTM [9]	80.1	0.77	1200	48
OASW Framework [4]	83.5	0.81	950	42
Proposed AM-LSTM	88.7	0.85	620	47

DISCUSSION

This consistent progression suggests that the architectural innovations introduced in the AM-LSTM notably the integration of an attention mechanism contribute meaningfully to its enhanced discriminative ability and improved handling of class imbalance. By prioritizing salient temporal features, the model is able to extract more informative patterns from sequential data, thereby strengthening its predictive reliability across diverse conditions.

Moreover, the AM-LSTM's substantial reduction in Drift Adaptation Sample Latency further underscores its practical value in real-world environments characterized by frequent or unpredictable data distribution shifts. Rapid drift responsiveness is essential in time-sensitive applications such as intrusion detection, network monitoring, and edge intelligence systems, where delays in model adaptation can lead to decreased detection accuracy or increased operational risk. The AM-LSTM's capacity to adapt at a significantly lower latency therefore represents a notable operational advantage.

Collectively, these findings reinforce the AM-LSTM as a technically rigorous and practically applicable model. Its balanced integration of accuracy, adaptability, and computational efficiency makes it a robust solution for dynamic, high-velocity data streams. Consequently, the model is well-positioned for deployment in real-time computational settings that demand both precision and responsiveness.

CONCLUSION

This study combines drift detection, online adaptation and multimodal fusion into a single adaptive multimodal LSTM(AM-LSTM) framework. Experimental results on two IoT benchmarks demonstrates that it can maintain higher accuracy and faster drift when compared to both static and adaptive alternatives. One of AM-LSTM's key advantages is its dynamic fusion method which makes the model robust to missing or incomplete datasets.

Additionally, it provides flexibility without incurring a significant computational burden due to its selective retraining strategy. We conducted the evaluation using the benchmarked datasets with artificial drift instead of real deployment. Also, evaluation was also done with scalability to high-dimensional multimodal data (such as video + audio + sensors) which might require further optimization through quantitation or data pruning.

FUTURE WORK

Future works will include as part of or research the ability to extend AM-LSTM to federates and privacy preserving environment. We will also be preserving said environment where adaptation occurs across distributed devices without sharing data. We will also include the use of transformer-based techniques to enhance modality fusion and enable better long-range sequences modelling.

Finally, we observed that using AM-LSTM framework is a step to achieve a unified and adaptive multimodal learning for IoTs. The proposed framework learns continuously under drift and maintains computing efficiency at the same time. This has huge potentials for smart homes and also applications in healthcare and systems monitoring in industries that operates in dynamic environments.

REFERENCES

[1] W. L. Hsu, M. Qiao, H. Xu, C. Zhang, H. L. Liu, and Y. C. Shiau, "Smart city governance evaluation in the era of internet of things: An empirical analysis of jiangsu, china," *Sustain.*, vol. 13, no. 24, 2021, doi: 10.3390/su132413606.

[2] J. Li, H. Yu, Z. Zhang, X. Luo, and S. Xie, "Concept Drift Adaptation by Exploiting Drift Type," *ACM Trans. Knowl. Discov. Data*, vol. 18, no. 4, 2024, doi: 10.1145/3638777.

[3] C. Ferreira and G. Gonçalves, "Remaining Useful Life prediction and challenges: A literature review on the use of Machine Learning Methods," 2022. doi: 10.1016/j.jmsy.2022.05.010.

[4] L. Yang and A. Shami, "A Lightweight Concept Drift Detection and Adaptation Framework for IoT Data Streams," *IEEE Internet Things Mag.*, vol. 4, no. 2, 2021, doi: 10.1109/IOTM.0001.2100012.

[5] K. Sathupadi, S. Achar, S. Bhaskaran, N. F.- Sensors, and undefined 2024, "Edge-cloud synergy for AI-enhanced sensor network data: A real-time predictive maintenance framework," *mdpi.com.*

[6] H. Wang, X. Liu, X. Zhong, L. Chen, F. Liu, and W. Zhang, "Multimodal Online Federated Learning with Modality Missing in Internet of Things," *arxiv.orgH Wang, X Liu, X Zhong, L Chen, F Liu, W ZhangarXiv Prepr. arXiv2505.16138, 2025•arxiv.org*, Accessed: Sep. 09, 2025. [Online]. Available: https://arxiv.org/abs/2505.16138