Guest Editorial: Special Issue on Stream Learning

IN RECENT years, learning from streaming data, commonly known as stream learning, has enjoyed tremendous growth and shown a wealth of development at both the conceptual and application levels. Stream learning is highly visible in both the machine learning and data science fields and has become a hot new direction in research. Advancements in stream learning include learning with concept drift detection, that includes whether a drift has occurred; understanding where, when, and how a drift occurs; adaptation by actively or passively updating models; and online learning, active learning, incremental learning, and reinforcement learning in data streaming situations.

As we have seen, these advancements have demonstrated how stream learning technologies can support machine learning capabilities in dynamic systems and environments. We have also witnessed compelling evidence of how stream learning can be used to support real-time data monitoring, analytics, prediction, and decision-making. Considering these observations, it is instructive, vital, and timely to offer a unified view of the current trends in the fundamental and applied research pertaining to stream learning to both improve the impact of machine learning and data science and to form a broad forum for advancing this integral stream of research.

The aim of this Special Issue is to gather and report latest progresses in the fundamental principles, practical methodologies, efficient implementations, and salient applications of stream learning. In this issue, we will attempt to cover the issues associated with concept drift and online learning that can help us to understand the knowledge discovered from large data streams when distributions change unpredictably, along with the mechanisms that can assist with online learning in dynamic situations.

This Special Issue offers well-focused, high-quality publications that report on significant results on the topic of stream learning. We hope to make visible the methods and applications that see this field shine. A total of 98 submissions were received from approximately 15 countries. Ultimately, 21 innovative and high-quality papers were selected, which significantly contribute to new theories, methodologies, algorithms, and applications to deal with challenging problems in stream learning. We hope this Special Issue will raise awareness about the stream learning technologies that are working within the machine learning and data science communities to handle challenging issues in stream mining.

Digital Object Identifier 10.1109/TNNLS.2023.3304146

The 21 impressive papers accepted for this Special Issue can be classified into five groups.

The first group comprises six articles that report on new developments in the online and active learning on data streams.

In [A1], da Silva et al. introduce an adaptive model based on resonance theory for unsupervised and semi-supervised online learning. The model demonstrates improved accuracy and robustness to ordering effects.

In [A2], Li et al. propose a multilayer framework for online metric learning (OML) that captures the nonlinear similarities among instances. A Mahalanobis-based OML algorithm is also presented, which is based on a passive-aggressive strategy combined with a one-pass triplet construction. This article also provides a theoretical analysis to explain the learning process and theoretically guarantee the work.

In [A3], Liu et al. develop a novel online active broad learning approach. The effectiveness of the broad learning system in the framework of online active learning is also revealed and verified. A reasonable dynamic asymmetric query strategy is then designed to mitigate the negative effects of class imbalance. The advantage of the proposed approach is that it helps to control the evolutionary direction of the stream learner.

The article [A4], by Liu et al., is an interesting study on online active learning algorithms for binary and multiclass online classification with trapezoidal data streams where the feature space may expand over time. In the context of an ever-changing feature space, this article provides a theoretical analysis of the mistake bounds and, as such, yields better classification accuracy.

In [A5], Wu et al. investigate feature selection in streaming data through incremental Markov boundary learning. The Markov boundary is learned by analyzing the conditional dependencies/independencies in streaming data. Furthermore, it avoids the negative impact of invalid prior information.

The last article in this category [A6], by Mastelini et al., takes inspiration from the batch learning extra trees algorithm. The proposed techniques give rise to competitive error predictions with significantly reduced computational costs.

These six articles highlight new developments and creative work in the context of online learning and active learning with data streams.

The second group of articles, consisting of three articles, focuses on concept drift detection and its theory and application.

In [A7], Green et al. target a highly innovative and challenging issue in concept drift detection. The article presents a framework for organizing signal-processing and machine-learning techniques to provide adaptive classification and concept drift detection. An illustrative case study is provided to demonstrate the proposed framework for drift tracking.

In [A8], Stucchi et al. address the problem of detecting distribution changes in a novel batch-wise and multimodal setup. It presents a drift detection algorithm that uses a single histogram to model batch-wise multimodal stationary conditions. It also shows the potential applications of the proposed algorithm in stream learning.

In [A9], Xu et al. propose a two-way concept-cognitive learning method for enhancing the flexibility and evolutionary ability of 2WL for concept learning. An example analysis with TEXT and corresponding experiments demonstrate the flexibility and reduced time-consumption of the proposed method.

These articles represent exceptional presentations of how to detect concept drift by analyzing large numbers of data streams obtained in various situations. Thus, they make a significant contribution to the field of concept drift.

The third group of five articles report new research development in drift and domain adaptation.

The first article [A10] by Fedeli et al. focuses on supervised learning problems in an online nonstationary data stream setting. The article introduces a novel learner-agnostic algorithm for drift adaptation where the learner is efficiently retrained when drift is detected. This drift adaptation method incrementally estimates the joint probability density of the input and target of the incoming data. Moreover, as soon as drift is detected, the learner is retrained using importance-weighted empirical risk minimization.

The second article [A11], by Yi et al., examines a general two-stage framework. The framework can train the target model by first learning a domain-level model, plus it can then fine-tune that model at the component-level to construct a bipartite graph which finds the most relevant component in the source domain for each component in the target domain.

In [A12], Weng et al. aim to attack the problem of extreme label shortages in cross domain, multistream classification problems. The proposed solution, called the "Learning Streaming Process from Partial Ground Truths," is built upon a flexible deep clustering network that delivers improved performances with cross domain adaptation.

In [A13], Du et al. explore how to deploy cooperative policies for homogeneous agents. The authors present a novel method that employs a heterogeneous graph attention network to model the relationships between heterogeneous agents in data streaming situations.

In [A14], Ren et al. present a new adaptive stream learning scheme for few-shot streaming tasks with the contributions of tensor and meta-learning. This scheme is conducive to mitigating domain shift when a new task only has a few labeled samples.

This group of articles presents the most innovative developments and solutions in drift adaptation in big streaming data research.

The fourth group, which contains three articles, presents new developments in data stream mining.

In [A15], Zhou et al. focus on semi-supervised streaming data mining. To classify data streams and detect novel classes, these researchers propose an algorithm that can handle any degree of separation between a novel class and some known classes. Furthermore, the algorithm requires only limited labeled instances to build a model.

In [A16], Wang et al. deal with uncertainty in evolving data streams. It introduces member meta-continual learning with a neural process for estimating uncertainty. The approach comprises two levels of uncertainty estimations—local uncertainty and global uncertainty—which each represent an evolution function in dynamic environments.

In [A17], Li et al. present novel thresholds of interval dominance degree and interval overlap degree between interval values. Interval-valued dominance relations are then constructed from the two developed parameters.

These articles discover the hidden patterns, correlations, insights, and knowledge to be found in data streams.

The fifth group, containing four articles, focuses on applications of stream learning.

To forecast urban traffic, in [A18], Chen et al. present an adaptive transformer. This article highlights a bidirectional spatial–temporal adaptive transformer for accurate traffic forecasting in data streaming situations.

In [A19], Zhang et al. present an important application of fake review detection. This article proposes a fake review detection model that can continuously learn a prediction model from a constantly arriving stream of data. The experimental results fully endorse that the proposed model effectively detects fake reviews, especially deceptive reviews.

In [A20], Yu et al. present a new approach in understanding the health status of patients at different stages of their treatment. A pretrained health progression network is developed by fusing heterogeneous medical information to create HealthNet. The researchers show that this method can capture the progression patterns of a patient's health in fine-grained detail.

In [A21], Dong et al. propose an advanced and effective incremental 3-D object recognition network (InOR-Net) that is able to recognize new classes of 3-D objects while overcoming catastrophic forgetting with old classes. These researchers also propose a strategy of dual adaptive fairness compensation to overcome the forgetting associated with class imbalances.

These four articles show that streaming data are common in practice and that stream learning capability is crucial for the good performance of machine learning systems and applications.

In summary, the articles included in this Special Issue represent significant advances in the state-of-the-art of stream learning that extends to models, algorithms, methodologies, and applications. These contributions not only report recent significant developments but also highlight potential, growing research directions and future trends that stand to benefit researchers in both the machine learning and data science fields.

We would like to take this opportunity to express our sincere thanks to all authors who have submitted their work to this Special Issue. Our thanks also extend to the reviewers, whose

expertise and critical, yet constructive, comments have been indispensable in improving the quality of the articles in this Special Issue.

JIE LU, *Guest Editor*
Australian Artificial Intelligence Institute
University of Technology Sydney
Ultimo NSW 2007, Australia
e-mail: Jie.Lu@uts.edu.au

JOAO GAMA, *Guest Editor*
Faculty of Economics
University of Porto
4099-002 Porto, Portugal
e-mail: joao.jgama@gmail.com

XIN YAO, *Guest Editor*
Department of Computer Science and Engineering
Southern University of Science and Technology
Shenzhen 518055, China
e-mail: xiny@sustech.edu.cn

Leandro Minku, *Guest Editor*
School of Computer Science
University of Birmingham
B15 2TT Birmingham, U.K.
e-mail: l.l.minku@bham.ac.uk

APPENDIX: RELATED ARTICLES

[A1] L. E. B. da Silva, N. Rayapati, and D. C. Wunsch, "Incremental cluster validity index-guided online learning for performance and robustness to presentation order," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 10, pp. 6686–6700, Oct. 2023.

[A2] W. Li et al., "A multilayer framework for online metric learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 10, pp. 6701–6713, Oct. 2023.

[A3] Z. Liu, Y. Zhang, Z. Ding, and X. He, "An online active broad learning approach for real-time safety assessment of dynamic systems in nonstationary environments," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 10, pp. 6714–6724, Oct. 2023.

[A4] Y. Liu, X. Fan, W. Li, and Y. Gao, "Online passive-aggressive active learning for trapezoidal data streams," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 10, pp. 6725–6739, Oct. 2023.

[A5] X. Wu, B. Jiang, X. Wang, T. Ban, and H. Chen, "Feature selection in the data stream based on incremental Markov boundary learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 10, pp. 6740–6754, Oct. 2023.

[A6] S. M. Mastelini, F. K. Nakano, and C. Vens, "Online extra trees regressor," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 10, pp. 6755–6767, Oct. 2023.

[A7] D. H. Green, A. W. Langham, R. A. Agustin, D. W. Quinn, and S. B. Leeb, "Adaptation for automated drift detection in electromechanical machine monitoring," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 10, pp. 6768–6782, Oct. 2023.

[A8] D. Stucchi, L. Magri, D. Carrera, and G. Boracchi, "Multimodal batch-wise change detection," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 10, pp. 6783–6797, Oct. 2023.

[A9] W. Xu, D. Guo, J. Mi, Y. Qian, K. Zheng, and W. Ding, "Two-way concept-cognitive learning via concept movement viewpoint," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 10, pp. 6798–6812, Oct. 2023.

[A10] F. Fedeli, A. M. Metelli, F. Trovò, and M. Restelli, "IWDA: Importance weighting for drift adaptation in streaming supervised learning problems," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 10, pp. 6813–6823, Oct. 2023.

[A11] C. Yi et al., "Multicomponent adversarial domain adaptation: A general framework," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 10, pp. 6824–6838, Oct. 2023.

[A12] W. Weng et al., "Autonomous cross domain adaptation under extreme label scarcity," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 10, pp. 6839–6850, Oct. 2023.

[A13] W. Du, S. Ding, C. Zhang, and Z. Shi, "Multiagent reinforcement learning with heterogeneous graph attention network," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 10, pp. 6851–6860, Oct. 2023.

[A14] B. Ren, L. T. Yang, Q. Zhang, J. Feng, and X. Nie, "Tensor-empowered adaptive learning for few-shot streaming tasks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 10, pp. 6861–6871, Oct. 2023.

[A15] P. Zhou, N. Wang, S. Zhao, Y. Zhang, and X. Wu, "Difficult novel class detection in semisupervised streaming data," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 10, pp. 6872–6886, Oct. 2023.

[A16] X. Wang, L. Yao, X. Wang, H.-Y. Paik, and S. Wang, "Uncertainty estimation with neural processes for meta-continual learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 10, pp. 6887–6897, Oct. 2023.

[A17] W. Li, H. Zhou, W. Xu, X.-Z. Wang, and W. Pedrycz, "Interval dominance-based feature selection for interval-valued ordered data," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 10, pp. 6898–6912, Oct. 2023.

[A18] C. Chen, Y. Liu, L. Chen, and C. Zhang, "Bidirectional spatial–temporal adaptive transformer for urban traffic flow forecasting," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 10, pp. 6913–6925, Oct. 2023.

[A19] Z. Shunxiang, Z. Aoqiang, Z. Guangli, W. Zhongliang, and L. KuanChing, "Building fake review detection model based on sentiment intensity and PU learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 10, pp. 6926–6939, Oct. 2023.

[A20] F. Yu et al., "HealthNet: A health progression network via heterogeneous medical information fusion," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 10, pp. 6940–6954, Oct. 2023.

[A21] J. Dong et al., "InOR-Net: Incremental 3-D object recognition network for point cloud representation," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 10, pp. 6955–6967, Oct. 2023.