

An Over view of Stream Mining of a Large Amount of Big Data

Andrew Mathews*

Data Science Engineer, Micron Technology Pvt Ltd. USA

Editorial

The method of obtaining knowledge structures from continuous, rapid data sets is known as data stream mining (also known as stream learning). Computer network traffic, phone calls, ATM transactions, web searches, and sensor data are all examples of data streams. The task of extracting insights from the Internet of Things (IoT) has been identified as one of the most intriguing and important opportunities for both academics and industry. As the number of applications requiring such processing grows, advanced analysis of huge data streams from sensors and devices is poised to become a prominent field of data mining research. One of the most challenging aspects of IoT stream mining is dealing with unbalanced data that evolves over time, i.e., concepts that drift or change totally. The goal of this collection is to examine the topic of learning from IoT data streams created by non-stationary processes that are developing. It will concentrate on the latest developments in distributed algorithms, approaches, methods, and tools for managing, exploiting, and interpreting data streams in non-stationary and imbalanced contexts. It will concentrate on modelling, prediction, and classification challenges based on data streams. The volume of data is rapidly increasing due to the development of the technology of information and communication. This data comes mostly in the form of streams. Learning from this ever-growing amount of data requires flexible learning models that self-adapt over time. In addition, these models must consider many constraints: (pseudo) real-time processing, high-velocity, and dynamic multi-form change such as concept drift and novelty. Consequently, learning from streams of evolving and unbounded data

requires developing new algorithms and methods able to learn under the following constraints: Random access to observations is either impossible or prohibitively expensive, memory is limited in comparison to the size of the data, and the data distribution or events that generate the data may change with time, a phenomenon known as idea drift. Over time, the number of classes may change. As a result, efficient data stream processing necessitates the following drivers and learning techniques: In order to incorporate the knowledge carried by each new arriving input, incremental learning is used. In order to forget or unlearn data samples that are no longer useful, incremental learning is used; novelty detection is used to learn new concepts. It's worth noting that streams are frequently generated by distributed sources, particularly with the rise of the Internet of Things. Learning algorithms that are scalable and decentralised have the potential to be more suited and efficient.

This collection provides a venue for researchers and practitioners to discuss key research problems and practical challenges in IoT big data stream mining. We seek submissions that describe and evaluate new (distributed) streaming learning algorithms, debates about standards, methodologies, tasks, and evaluation, open challenges, visualisation, applications, and comparisons of rival approaches. Alternative opinions should be represented, and conversations should be encouraged. We seek papers reporting cutting-edge research on all aspects of IoT big data stream mining to be submitted. Theoretical foundations, algorithms, systems, distributed and parallel approaches, federate learning, applications, language difficulties, data storage, access, and architecture are all encouraged in the papers.

*Address for Correspondence: Andrew Mathews, Data Science Engineer, Micron Technology Pvt Ltd. USA, E-mail: andrewmathew@gmail.com

Copyright: © 2021 Mathews A. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Received 21 October 2021; **Accepted** 26 October 2021; **Published** 02 November 2021

How to cite this article: Andrew Mathews. "An Over view of Stream Mining of a Large Amount of Big Data." *Ind Eng Manage* 10 (2021): 318.