Call For Papers

Special Issue:
"Data Stream Mining and Applications"

Co-Organizers

Edwin Lughofer, University of Linz, Austria, edwin.lughofer@jku.at
Robi Polikar, Rowan University, Glassboro, New Jersey, United States, polikar@rowan.edu
Mu-Yen Chen, Department of Information Management, National Taichung University of Science and Technology, Taiwan, E-mail: mychen@nutc.edu.tw

Scope

In current industrial systems, the necessity of ***data stream mining and learning from data streams*** is increasingly becoming more prevalent and urgent, due to speed, volume and on-line nature of the data generated by such systems. While conventional batch and off-line training approaches provide a possible solution, such approaches are often too time and memory intensive, and cannot process the data at the high enough rate that is often desired. This is true even when batch and off-line approaches are applied to sliding windows or onto streaming samples gathered from reservoir computing techniques.

An important aspect in data stream mining is that the data analysis system, the learner, has no control over the order of samples that arrive over time --- they simply arrive in the same order they are acquired and recorded. Also, the learning algorithms usually have to be fast enough in order to cope with (near) real-time and on-line demands. This usually requires a single-pass learning procedure, restricting the algorithm to update models and statistical information in a sample-wise manner, without using any prior data (or at least to aim for using as little prior data as possible). In literature, this is also termed as *incremental or sequential learning* and plays a key role in data stream mining frameworks and environments.

In some cases, it is sufficient to only update (some of) the model parameters, whereas in other cases the evolution of new structural components (e.g. adding a new neuron, rule or leaf) may be necessary in order to expand models on-the-fly and on demand to new operating modes, dynamically changing states or non-stationary environmental conditions.

Specific requirements for incremental/evolving learning mechanisms include:

1.) ability to track any data distributions that are common in streaming environments and to handle noise adequately for a robust process and highly stable models;
2.) ability to learn models from scratch with little parameterization effort as possible – ideally plug-and-play functionality is embedded (either no learning parameters or learning parameters which self-adapt over time).
3.) ability to process data stream samples in a very fast manner (due to on-line and real-time system demands), usually requiring small computational footprint and single pass functionality.
4.) ability to learn in a semi-supervised or even unsupervised manner, if and when the target values are costly or difficult to obtain;
5.) ability to learn from high dimensional data and/or built-in mechanisms for variable selection and dimensionality reduction;
6.) ability to track a dynamically changing or drifting environment in (near) real- time (fast detection, reaction and recovery).

This special issue intends to draw a picture of the recent advances in data stream mining techniques including all incremental machine learning concepts and evolving soft computing modeling strategies for addressing these important problems discussed above.

Particularly, the special issue aims at soliciting contributions dealing with real-world applications that present dynamic facets requiring on-line learning capabilities (see Topics below). Contributions on ***any form of evolving models/learners applied to any advanced data mining concepts in a stream learning context*** such as active learning, dynamic feature weighting/selection, drift analysis in data streams, incremental optimization of parameters, complexity reduction issues, dynamic ensemble and data fusion systems, outlier treatment as well as reliability issues are of high relevance.

Especially, dynamic feature weighting/selection, outlier treatment (to distinguish between a new operation mode and an outlier case), incremental optimization (to adapt parameters based on a mathematically formulated optimization function) as well as reliability issues pointing to the degree of certainty in model predictions have been underexplored in literature so far, and thus are given a special emphasis within this special issue. Furthermore, the approaches dealing with plug-and-play functionality are quite rare in literature as all stream modeling algorithms and evolving systems techniques usually possess several free parameters that require a manual fine-tuning, which can become a real handicap in real on-line learning environments. Also, hybrid methods for refining knowledge-based expert models with data streams over time to maintain high quality have been hardly handled in literature so far.

Finally, all new emerging and grand-challenge topics such as interpretability aspects in evolving models, cognitive interaction modes (active learning and teaching) as well as mimicking intelligent brain – even if at a limited scale - are of particular interest to this special issue. Computational aspects such as real-time capability of the learning methods play a central role within all approaches.

Topics

All contributions relevant to learning from data streams, in particular those associated with non-stationary or changing distributions are welcome!

We provide a list of *a few* core themes as valuable examples for the special issue:

- **Data Stream Mining with Soft Computing Techniques such as *(but not nec. restricted to)***:
    - Evolving fuzzy systems (EFS) and classifiers (EFC)
    - Sequential radial basis functions networks
    - Sequential multilayer perceptron
    - Online probabilistic neural networks
    - Evolving bio-inspired approaches
    - On-line genetic-based modelling systems and dynamic evolutionary algorithms
    - Any form of online, evolving hybrid (e.g. neuro-fuzzy, neuro-genetic) approaches

- **Data Stream Mining with Machine Learning and Data Mining Concepts such as *(but not necessarily restr. to)*:**
    - Online and incremental support vector machines
    - Evolving self-organizing maps
    - Incremental decision trees
    - Incremental ensemble classifier and trained fusion techniques
    - Bagging and Boosting in data streams
    - Evolving cluster models and incremental unsupervised learning
    - Incremental statistical learning techniques in non-stationary environments
    - Incremental Kernel-based learning and density ratio matching

- **Advanced Aspects for Improved Stability and Useability *(but not necessarily restr. to)***:
    - Concepts to address *drifts and shifts* in Data Streams
    - Concepts to address domain adaptation
    - Concepts to address importance weighting and sampling
    - On-line single-pass active learning from Data Streams
    - Semi-supervised learning from Data Streams
    - Fast adaptive, incremental learning methods ("hazard" form)
    - Dynamic dimension reduction and feature selection in Data Streams
    - Reliability in model predictions and parameters
    - Hybrid modelling aspects (refining knowledge-based models with data)
    - Parameter-low and –insensitive learning methods
    - On-line complexity reduction to emphasize transparent, more compact models
    - Concepts to address linguistic interpretability
    - Concepts to address visual interpretability (model development over time)
    - Online tuning via human-machine interaction

- **Real-World Applications of data stream mining such as *(but not necessarily restricted to)***:
    - Data stream modelling and identification (supervised and unsupervised)
    - *Big Data* handled in a streaming concept
    - Online fault detection and decision support systems
    - Online media stream classification
    - Process control and condition monitoring
    - Modeling in high throughput production systems
    - Web applications
    - Adaptive chemometric models in dynamic chemical processes

- Online time series analysis and stock market forecasting
- Robotics, Intelligent Transport and Advanced Manufacturing
- Adaptive Evolving Controller Design
- User Activities Recognition
- Cloud Computing
- Multiple Sensor Networks

Important dates

Submission deadline (EXTENDED):	**31^{st} of October, 2015**
First author notification:	15^{th} of January, 2016
Revised version:	15^{th} of March, 2016
Final notification:	31^{th} of May, 2016
Publication:	Summer 2016

The papers should be formatted according the instruction guidelines for authors found at
http://www.springer.com/physics/complexity/journal/12530
and submitted through the regular submission gate at @ Evolving Systems journal (Springer) located
http://www.springer.com/physics/complexity/journal/12530
by choosing ***S.I.: Data Stream Mining and Applications*** as article type.