

WSOM+ 2019 special issue

Editorial: A self-organizing world. Special issue of the 13th edition of the Workshop on Self-Organizing Maps and Learning Vector Quantization, Clustering and Data Visualization, WSOM+ 2019.

Alfredo Vellido (corr. author), Cecilio Angulo and Karina Gibert

The vibrant Mediterranean city of Barcelona, in Spain, welcomed the 13th edition of WSOM+ 2019, the Workshop on Self-Organizing Maps and Learning Vector Quantization, Clustering and Data Visualization. It was hosted by the Intelligent Data Science and Artificial Intelligence (IDEAI-UPC) Research Center at Universitat Politècnica de Catalunya (UPC BarcelonaTech) and sponsored by several companies, including Telefónica Alpha (now Kora Health), Amalfi Analytics, LumenAI and PredictAssistance. This was the latest of a series of biannual international conferences that started as WSOM in Helsinki, Finland, back in 1997, under the intellectual inspiration of Prof. Teuvo Kohonen, with a focus on Self-Organizing Maps. In its 12th edition, it extended its scope to cover the more general domains of unsupervised learning and data visualization.

WSOM+ 2019 included talks by renowned keynote speakers including professors Paulo J.G. Lisboa from Liverpool John Moores University, U.K. (Unsupervised ML in Sports Analytics), Tobias Schreck, from TU Graz, Austria (Visual Analytics Techniques for Data Exploration: Visual Cluster Analysis, Interactive Data Modeling, and User Guidance), Aïda Valls, from Universitat Rovira i Virgili, Spain (Ontology-based Clustering for Managing Semantic Data and its Use in Tourism Applications) and Alessandro Sperduti, from Università degli Studi di Padova, Italy (Self-organizing Maps for Structured Data: from Unsupervised to Deep Supervised Models). The variety of these talks reflected the own variety of the advances presented by the papers accepted in the conference.

The conference was structured in nine sessions on seven different topics: SOM theory, methods and practical applications covered three sessions; correspondingly, LVQ theory, methods and practical applications covered two sessions; then, two sessions were devoted to general theory and methods and general practical applications and, finally, two sessions to specific applications to the life sciences. The latter bear witness to the rapidly increasing interest in machine learning in general and unsupervised machine learning in particular shown by biology and medicine, mostly driven by developments in the omics sciences and the area of systems biology. A total of thirty-three papers were accepted for oral presentation. Out of these, ten were selected for extension and these are the ones that we now proudly present to the Neural Computing & Applications readership in this special issue.

The four first papers of this issue are devoted to SOM models. The first of them [*ref paper#1*] stems from one of the invited talks and merges self-organization and deep learning. Pasa, Navarin and Sperduti propose a SOM-based aggregation approach for Graph Convolutional Neural Networks. Deep learning of graphs is a hot topic due to the increasing availability of data that are naturally represented in a graph form (life sciences, social networks, etc.). A main component of a graph neural network is the aggregation operator required to generate a graph-level representation from a set of node-level embeddings. The proposed SOM-based *DeepSets* aggregation operator allows to compute node representations that embed information about their mutual similarity. The second SOM-related paper [*ref paper#2*] also addresses graphs, but from a very different standpoint. Taylor and Merenyi present DM-Prune: not a clustering method as such, but as a sparsification step for cumulative adjacency (CADJ)

graphs as inputs to graph-based clustering (GBC) methods for partitioning the learned prototypes of a vector quantizer. Even if CADJ graphs may originate from the prototypes of any vector quantizer, authors focus on prototypes obtained using SOM. The method is illustrated with GBC of images. The third paper [**ref paper#3**] by Ma, Kirby and Peterson takes a more theoretical road and describes the mathematical framework for implementing SOM mappings on flag manifolds. A flag is described as a nested sequence of vector spaces and a flag manifold as a manifold whose points parameterize all flags of a fixed type in a fixed vector space. Flags are argued to arise in data analysis contexts such as wavelet, Fourier, and singular value decompositions. Finally, in [**ref paper#4**], Bacciu, Bertoni and Morelli address an eminently practical problem: quality inspection and intelligent filtering of smart-bracelet data in the context of preventive analysis of heart activity sensors, plagued by noise issues introduced by motion artifacts. SOM models are shown to be effective for filtering sections of heart data collected from unreliable devices, such as wrist-worn devices.

The next four papers of the WSOM+ 2019 special issue deal with LVQ models, again from different perspectives. The first [**ref paper#5**] draws us towards a current -at the time of publication- affair: the COVID-19 pandemic. Kaden and co-workers investigate SARS-CoV-2 virus sequences based on alignment-free dissimilarity measures for RNA sequence comparison leading to virus type classification using an LVQ variant. The second LVQ-related paper [**ref paper#6**] proposes a quantum-inspired LVQ for prototype-based classification. In their paper, Villmann and co-workers are interested in interpretable variants such as Generalized LVQ (GLVQ) and propose a quantum computing-inspired realization of this model. Kernelized GLVQ methods are described and linked to quantum space ideas using quantum bit vector space in the quantum state space. The third paper [**ref paper#7**] takes the reader to another timely topic: concept drift handling, with a focus on streaming data. Heusinger, Raab and Schleif modify Soft LVQ and GLVQ to handle streaming data. Here, momentum-based stochastic gradient descent techniques are applied to tackle concept drift passively. The fourth and last paper of this batch [**ref paper#8**] also addresses the problem of concept drift. Straat and co-workers apply methods from statistical physics to develop a modelling framework for the mathematical analysis of the training dynamics in non-stationary environments, showing that standard LVQ algorithms are, to some extent, suitable for the training in non-stationary environments without modification. Authors go beyond LVQ to investigate gradient-based training of layered neural networks with sigmoidal activation functions, comparing them with the use of rectified linear units (ReLU), showing that the sensitivity to concept drift differs significantly between these types of activation function.

The remaining two papers in the issue tackle quite different problems. The work by Karna and Gibert [**ref paper#9**] provides solutions for the classical problem of determining the most adequate number of clusters in hierarchical clustering settings. Usually, this determination from the data partition dendrogram requires human expertise as input, especially in real application settings like the one which is the focus of the paper: decision support in an industry 4.0 environment, namely 3D printers based on sensor data in real-time. The final paper [**ref paper#10**] takes us back to the problem of concept drift. Here, Szadkowski addresses, in a combined manner, two challenges for life-long classifiers: concept drift and catastrophic forgetting. For that, the process of continual learning is defined as a wrap modification, where a wrap is a manifold that can be trained to cover or uncover a given set of samples and is used as a theoretical building block for a general life-long learning scheme implemented as an ensemble of variational autoencoders.

The chairpersons of WSOM+ 2019, as editors and curators of this special issue, hope that you will enjoy reading the collected contributions and find them useful. We encourage you to submit your work to, and follow the future editions of, our WSOM+ conference. We express our gratitude to the editor-in-chief for his invitation to produce this issue of the journal. We also thank the editorial staff for their kind help on its elaboration.

Alfredo Vellido, on behalf of the chairing committee (Alfredo Vellido, Cecilio Angulo and Karina Gibert).

[*ref paper#1*] NCAA-D-20-01350R1

[*ref paper#2*] NCAA-D-20-01365R1

[*ref paper#3*] NCAA-D-20-01316R1

[*ref paper#4*] NCAA-D-20-01298R1

[*ref paper#5*] NCAA-D-20-01991R2

[*ref paper#6*] NCAA-D-20-01392R2

[*ref paper#7*] NCAA-D-20-00542R1

[*ref paper#8*] NCAA-D-20-01477R1

[*ref paper#9*] NCAA-D-20-02157R2

[*ref paper#10*] NCAA-D-20-01231R2