Guest Editorial for Special Issue on Time Series Classification

Houshang Darabi, *Senior Member, IEEE*, Georgiana Ifrim, Patrick Schäfer, and Diego Furtado Silva

TIME series classification research and applications have recently received a lot of attention. With the rapid advancement of technological devices, time series data are being collected by a wide variety of devices, resulting in a wide range of research and applications. Various fields of studies, ranging from healthcare to weather readings, require time series classification. Activity and action recognition using time series data from fitness trackers are commonly being practiced. Today, a new generation of algorithms and techniques are being used to classify time series data that are extensively available to the research community. For example, recently deep learning techniques that utilize convolutions and recurrent neural networks are being used to classify epileptic seizures from EEG recordings made available at the University of California-Irvine data repository.

This special issue received 17 submissions. Four of these submissions were selected for publication in the special issue (an acceptance rate of 23%). A total of eight papers appeared in the special issue of which four were not special issue submissions. These four papers also underwent a rigorous review process conducted by the editorial board of the journal. The papers selected for the special issue cover a range of research problems in the areas of time series classification and machine learning. They report on both the theory and application advances made in these areas. While each one of these papers has its own merits and is of interest to the research community, we would like to encourage the readers to pay special attention to the paper titled "The UCR Time Series Archive". The (UCR) times series archive is the most recognized resource by the time series classification community and it has been extensively cited by the members of the community. This specific paper discusses the recent extensions made to the archive. It also provides "advice to the community on best practices on using the archive to test classification algorithms". We hope the time series classification research community, and especially its new researchers, find this paper useful. The other accepted papers are introduced below.

The paper "Classification of Short Time Series in Early Parkinson's Disease with Deep Learning of Fuzzy Recurrence Plots" is authored by Tuan Pham, Karin Anders Eklund, Wårdell, and Göran Salerud from the Linköping University. This paper introduces the use of fuzzy recurrence plots of very short time series, as input data for the training and classification with long short-term memory neural networks. The authors show that fuzzy recurrence plots provide promising results and outperform the direct input of the time series for the classification of healthy control and early Parkinson's Disease subjects.

The paper "Self-Learning of Multivariate Time Series Using Perceptually Important Points" is authored by Timo Lintonen and Tomi Räty from the VTT Technical Research Centre of Finland, Finland. This paper proposes a novel stopping criterion for self-learning with time series data, called peak evaluation using perceptually important points. The authors have shown that their Peak evaluation method performed well on both univariate and multivariate time series classification datasets, with the advantage that it is easy to use since it does not require setting any hyperparameters.

The paper "Clustering Structure Analysis in Time-Series Data With Density-Based Clusterability Measure" is authored by Juho Jokinen, Tomi Räty and Timo Lintonen from the VTT Technical Research Centre of Finland, Finland. This paper proposes a novel clusterability assessment method called density-based clusterability measure, designed to avoid discovering false structures in time series data. The performance of the new clusterability measure is evaluated against several synthetic data sets and time-series data sets and shows that the density-based clusterability measure can successfully indicate clustering structure of time-series data.

The paper "Long-term Traffic Volume Prediction Based on K-means Gaussian Interval Type-2 Fuzzy Sets" is authored by Runmei Li, Yinfeng Huang, and Jian Wang from Beijing Jiaotong University, China. This paper uses Gaussian interval type-2 fuzzy set theory on historical traffic volume data processing to obtain a 24-hour prediction of traffic volume with high precision. The simulation results show that the proposed forecasting method achieves a much lower average relative error than existing approaches.

The paper "Efficient Deviation Detection Between a Process Model and Event Logs" is authored by Lu Wang, Yuyue Du, and Liang Qi from Shandong University of Science and Technology, China. This paper focuses on business processes and proposes an efficient approach to

Citation: H. S. Darabi, G. Ifrim, P. Schäfer, and D. Silva, "Time series classification," *IEEE/CAA J. Autom. Sinica*, vol. 6, no. 6, pp. 1291–1292, Nov. 2019.

Houshang Darabi is with University of Illinois at Chicago, USA (e-mail: hdarabi@uic.edu).

Georgiana Ifrim is with University College Dublin, Ireland (e-mail: georgiana.ifrim@insight-centre.org).

Patrick Schäfer is with Humboldt University of Berlin, Germany (e-mail: patrick.schaefer@hu-berlin.de).

Diego Silva is with Federal University of São Carlos, Brazil (e-mail: diegofs@ufscar.br).

Digital Object Identifier 10.1109/JAS.2019.1911741

detect deviations between a process model and event logs. A clinical process in a healthcare information system is used as a case study to illustrate the effectiveness of the proposed approach.

The paper "Forecasting of Software Reliability Using Neighborhood Fuzzy Particle Swarm Optimization Based Novel Neural Network" is authored by Pratik Roy and Kashi Nath Dey from University of Calcutta, India and Ghanshaym Singha Mahapatra from National Institute of Technology, India. This paper proposes an artificial neural network based software reliability model trained by a novel particle swarm optimization algorithm for enhanced forecasting of the reliability of software. The authors show that faster release of software is achievable by applying the proposed PSO based neural network model during the testing period.

The paper "Predicting the Results of RNA Molecular Specific Hybridization Using Machine Learning" is authored by Weijun Zhu, Xiaokai Liu and Mingliang Xu from Zhengzhou University, China and Huanmei Wu from Indiana University-Purdue University Indianapolis, USA. This paper introduces a machine learning based technique to decide within an acceptable time, whether a specific RNA hybridization is effective. Given an RNA design, a Boosted Tree-based approach demonstrates high computational efficiency and better predictive accuracy in determining the biological effectiveness of molecular hybridization.

We would like to thank all the anonymous reviewers who performed timely and high quality reviews of the submitted articles to this special issue. Our special thanks go to the Editor-in-Chief, Professor Mengchu Zhou, whose valuable advice was instrumental in the successful completion of this special issue. Finally, we would like to express our gratitude to Dr. Yan Ou, the Managing Editor of the journal, who was always there for us and helped us with all the logistics and clerical tasks related to this special issue.

Houshang Darabi (S'98–A'00–M'10–SM'14) received the Ph.D. degree in industrial and systems engineering from Rutgers University, New Brunswick, NJ, USA, in 2000.

He is currently an Associate Professor with the Department of Mechanical and Industrial Engineering, University of Illinois at Chicago (UIC), and also an Associate Professor with the Department of Computer Science, UIC. His research has been supported by several agencies, such as the National Science Foundation, the National Institute of Standard and Technology, and the Department of Energy. He has extensively published on various subjects, including time series classification, and process mining. His current research interests include the application of data mining, process mining, and optimization in design and analysis of manufacturing and healthcare systems.

Georgiana Ifrim holds a Ph.D. and M.Sc. degree from Max-Planck Institute for Informatics, Germany, and a B.Sc. degree from University of Bucharest, Romania.

She is an Assistant Professor at the School of Computer Science, University College Dublin, Ireland, Co-Lead of the SFI Centre for Research Training in Machine Learning (ML-Labs) and SFI Funded Investigator with the Insight Centre for Data Analytics and VistaMilk research centres. Dr. Ifrim's research focuses on developing scalable predictive models for machine learning and data mining applications. She has developed new methods for sequence learning, time series classification, text mining and real-time prediction for news and social streams. Her current research focuses on the design of efficient and interpretable learning models for sequences (e.g., DNA, time series), and on real-time prediction for streaming data (e.g., news and social media).

Patrick Schäfer holds a Ph.D. degree from Humboldt University of Berlin and a M.Sc. degree from Free University of Berlin, both in Computer Science.

He is a Postdoc Researcher and Lecturer at the Humboldt University of Berlin. He Besides he worked at the Konrad Zuse Institute in Parallel and Distributed Systems in Berlin. His main research interests are scalable time series analytics and parallel and distributed systems. His current research is on early and scalable time series classification, time series motif discovery, and assessing the land use using satellite image time series.

Diego Furtado Silva holds a M.Sc. and Ph.D. degrees in Computer Science and Computational Mathematics at the Institute of Mathematics and Computer Sciences (ICMC), University of São Paulo (USP), where he also graduated in Computer Science.

He is an Assistant Professor at the Federal University of São Carlos. Besides, he has worked at the University of Columbia and the University of California, Riverside. His main research interests are time series mining, music information retrieval, and data stream classification. His current research approaches these domains using, primarily, meta-learning and deep learning techniques.