Building Adaptive Data Mining Models on Streaming Data in Real-Time, an Outlook on Challenges, Approaches and Ongoing Research

Frederic Theodor Stahl

University of Reading
Department of Computer Science
f.t.stahl@reading.ac.uk

ABSTRACT

Advances in hardware and software, in the past two decades have enabled the capturing, recording and processing of potentially large and infinite streaming data. As a consequence the field of research in Data Stream Mining has emerged building Data Mining models, workflows and algorithms enabling the efficient and effective analysis of such streaming data at a large scale. Application areas of Data Stream Mining techniques include real-time telecommunication data, telemetric data from large industrial plants, credit card transactions, social media data, Smart Cities, IoT, etc. Some applications allow the data to be processed modelled and analysed in batches by traditional Data Mining approaches. However, others require the model building and analytics to take place in real-time as soon as new data becomes available i.e. to accommodate infinite streams and fast changing concepts in the data. This talk discusses challenges, barriers, opportunities and recent research on Micro-Cluster based Data Stream Mining models to overcome these barriers.

Biography

Frederic Stahl is Associate Professor in Data Science at the University of Reading and has been working in the field of Data Mining and Knowledge Discovery from Data (KDD) for the last 12 years. In particular he has been working in the research domain

Proceedings 32nd European Conference on Modelling and Simulation ©ECMS Lars Nolle, Alexandra Burger, Christoph Tholen, Jens Werner, Jens Wellhausen (Editors)
ISBN: 978-0-9932440-6-3/ ISBN: 978-0-9932440-7-0 (CD)

of Big Data Analytics. His research interests here are in (i) developing scalable parallel algorithms for building Data Mining models on large volumes of data; (ii) developing algorithms for building self-adaptive Data Mining models for real-time streaming data and (iii) applications in Big Data Analytics. He currently leads a small group of 5 PhD students working on many aspects of Data Mining and Data Stream Mining. In previous appointments Frederic worked as a Lecturer at Bournemouth University and as Senior Research Associate at the University of Portsmouth. He obtained his PhD in 2010 from the University of Portsmouth with the title “Parallel Rule Induction” and his Engineering Diploma in Bioninformatics in 2006 from the University of Applied Science Weihenstephan (Germany). He has published over 50 articles in peer-reviewed conferences and journals. He is heavily involved in the BCS SGAI, the Specialist Group on Artificial Intelligence of the British Computer Society. Here he serves as an elected committee member, is the main organiser of the UK Symposium on Knowledge Discovery and Data Mining, co-organiser of the societie’s annual International Conference on Artificial Intelligence and Guest Editor of the conference’s Special Issue journal (Expert Systems, Wiley).