Data Preparation for Data Mining

Special Track for the Ninth International Conference on Computer Science &
Software Engineering – C3S2E'16 (http://confsys.encs.concordia.ca/C3S2E/c3s2e16/)
July 20 – 22, 2016, Porto, Portugal

Call for Papers

Submission deadline: March 20, 2016

Context

Current technological developments allow the collection of huge amounts of data that can be used to support decision-making processes. However, this is only possible if data can be transformed into knowledge.Variety of data mining algorithms are used to extract data patterns. Tasks for pattern extraction include classification (rules or trees), regression, clustering, association, sequence modeling, dependency, and so forth. However, much work in the field of data mining was built on the existence of data with quality, and real-world data is often incomplete, noisy, or inconsistent, representing an obstacle for efficient data analysis and mining. Other challenges include big data (efficiency, parallel processing and the number of features and examples), curse of dimensionality, or the use of domain knowledge. Although most mining algorithms have some procedures for dealing with dirty data, they lack for robustness. Furthermore, low-quality data will lead to low-quality data analysis and mining results (Garbage in, garbage out). Data preparation techniques, when applied before mining, can substantially improve the overall quality of the data and consequently improve the mining results and/or the time required for the mining process. Thus, the development of data preparation techniques is both a challenging and a critical task. This Special Track on Data Preparation for Data Mining will address practical techniques and methodologies of data preparation for data-mining applications.

Topics of interest
Topics include but are not limited to:
-Data collecting
-Data integration
-Data reduction
-Data cleaning
-Detection of outliers
-Data/Information quality
-Data profiling
-Data enrichment
-Feature selection and transformation
-Data summarization
-Data discretization
-Data encoding
-Sampling
-Data preparation on regression/classification
-Data preparation on segmentation/clustering
-Data preparation on association rules

-Data preparation on text mining
-Data preparation on web mining
-Data preparation on visual data mining
-Data preparation on temporal and spatial data mining
-Data preparation on multimedia mining (audio/video)

Submission instructions

Prospective authors interested in participating and presenting work in this Special Track are invited to submit a contribution (either a full or short paper). Both theoretical and practical contributions are encouraged.

- Full papers - original, relevant and previously unpublished research results, related to any of the topics of the Special Track, with a maximum of 10 pages.
- Short papers - project reports, research in progress, with no more than 4 pages.

Accepted papers will be presented and discussed in time slots of 25 (full papers) or 15 minutes (short papers).

Papers must be written in English. Submissions should follow the guidelines of the C3S2E'16, namely the types of papers (full or short) and formatting rules, available in the following link: http://confsys.encs.concordia.ca/C3S2E/c3s2e16/c3s2e16_author_kit.php.

All papers should be submitted using the ConfSys system, available in the following link: http://confsys.encs.concordia.ca/ConfSys/jsp/index.jsp.

Publication

Accepted papers will be published in the conference proceedings. The conference proceedings will be published by ICPS (ACM) in their digital library; the ISBN assigned by ACM to C3S2E 2016 is: 978-1-4503-4075-5. A version of the proceedings to be distributed to the conference attendees would be prepared by BytePress.

Important dates
Papers submission deadline: March 20, 2016
Notification of acceptance: May 23, 2016
Camera-ready deadline: June 24, 2016
Conference: July 20-22, 2016

Organizing Committee:
- Pedro Henriques, Universidade do Minho, pedrorangelhenriques@gmail.com
- Fátima Rodrigues, Instituto Superior de Engenharia do Porto, mfc@isep.ipp.pt
- Paulo Oliveira, Instituto Superior de Engenharia do Porto, pjo@isep.ipp.pt
- Alberto Freitas, Faculdade de Medicina da Universidade do Porto, alberto@med.up.pt