CALL FOR RESEARCH PAPERS

MultiClust: Discovering, Summarizing and Using Multiple Clusterings
Special Issue of the Machine Learning journal

Editors: Emmanuel Müller, Ira Assent, Stephan Günnemann, Thomas Seidl and Jennifer Dy

Traditionally, clustering has focused on determining a single summary of the data. In today's applications, however, data is collected for multiple analysis tasks. Several features or measurements provide complex and high dimensional information. In such data, one typically observes several valid groupings, i.e. each data object fits in different roles. In contrast to traditional clustering these alternative clusterings describe multiple aspects that characterize the data in different ways.

The topic of multiple clustering solutions by itself shows multiple research aspects: multiple alternative solutions vs. a single consensus that integrates different views; given views in multi-source clustering vs. detection of novel views by feature selection and space transformation techniques; a virtually unlimited number of alternative solutions vs. a non-redundant output restricted to a small number of disparate clusterings. Further aspects are induced by complex data types ranging from heterogeneous vector spaces with continuous, categorical, and binary data up to data representations using graphs, sequences, streams, etc.

The topic of multiple clustering solutions has opened novel challenges in a number of research fields. Examples from the machine learning and knowledge discovery communities include frequent itemset mining, ensemble mining, constraint-based mining, theory on summarization of results, or consensus mining to name only a few. We observe fruitful input from these established related areas. Overall, this cross-disciplinary research endeavor has recently received significant attention from multiple communities.

The aim of this special issue is to establish an overview of recent work for the growing community interested in multiple clustering solutions. It intends to increase the visibility of the topic itself but also to bridge it to closely related research areas. As a platform for research in a young field, the journal is open to submissions from both newcomers and experts in this area.
TOPICS OF INTEREST

This special issue focuses on the emerging topic of “Multiple Clustering Solutions”. It is addressed to the emerging communities in Database Systems, Machine Learning, Data Mining, and Statistics that are contributing to this topic. Recent MultiClust workshops at KDD 2010, ECML PKDD 2011, and SDM 2012 and a recent tutorial on “Discovering Multiple Clustering Solutions” (at ICDM 2010, SDM 2011, and ICDE 2012) document the research interest in this exciting topic.

A non-exhaustive list of topics of interest is given below:

- Discovering multiple clustering solutions
  - Alternative / disparate / orthogonal clusters
  - Multi-view clustering / subspace clustering / co-clustering
  - Multi-source clustering / clustering in parallel universes
  - Feature selection and space transformation techniques
  - Constraint-based mining for the detection of alternatives
  - Non-redundant view and cluster detection
  - Model selection problem:
    - how many clusterings / how many clusters
  - Iterative vs. simultaneous processing of multiple views
  - Scalability to large and high dimensional databases
  - Targeting complex data types
    - (e.g. graphs, sequences, or streams)

- Summarizing multiple clustering solutions
  - Ensemble techniques
  - Meta clustering
  - Consensus mining
  - Summarization and compression theory

- Using and evaluating multiple clustering solutions
  - Classification based on multiple clusterings
  - Evaluation metrics for multiple clustering solutions
  - Visualization and exploration of multiple clusterings

- Related research fields
  - Frequent itemset mining
  - Subgroup mining
  - Subspace learning
  - Relational data mining
  - Transfer mining

- Applications of multiple clustering solutions
  - Bioinformatics: gene expression analysis / proteomics / ...
  - Sensor network analysis
  - Social network analysis
  - Health surveillance
  - Customer segmentation
  - ... and many more ...

We encourage also submissions describing innovative work in other related fields that address the issue of multiplicity in clustering.
SUBMISSION GUIDELINES

Submissions to the special issue must be submitted like regular submissions to the journal. Instructions can be found at http://www.springer.com/computer/ai/journal/10994

Extended Submission Deadline: May 25, 2012
Choose Article Type: "SI: MultiClust"

If you are considering submitting to the special issue and have questions regarding the scope or need further information, please do not hesitate to contact the editors.

EDITORS

Emmanuel Müller, emmanuel.mueller@kit.edu
Karlsruhe Institute of Technology, Germany

Ira Assent, ira@cs.au.dk
Aarhus University, Denmark

Stephan Günnemann, guennemann@cs.rwth-aachen.de
RWTH Aachen University, Germany

Thomas Seidl, seidl@cs.rwth-aachen.de
RWTH Aachen University, Germany

Jennifer Dy, jdy@ece.neu.edu
Northeastern University, USA