Workshop on Tools and Technologies in Statistics, Machine Learning and Information Retrieval for Educational Data Mining
(SMLIR@EDM2015)

Within Intelligent Data Analysis (IDA) research area there may be found several subareas such as Machine Learning/Statistics, Information Retrieval, Data Mining and lately Big Data and Cloud Computing. The exact boundaries between these research areas are not very clear and may mislead research efforts, each one having its own particularities in terms of input type and size, data processing methodologies and obtained output. As Data Mining makes intensive use of all these research subareas, it is mandatory to be aware of their subtle differences and, therefore, design and implement IDA systems that make research efforts sound and effective. The goal of this workshop is to gather research efforts that fall into any of the categories of subareas and have results into the application area of Education. The workshop is looking for contributions that provide experimental results in the area of EDM/Information Retrieval and are focused on data processing fundamentals from Machine Learning/Statistics perspective. From the practical point of view, the focus should be on presenting the details regarding what and how tools and technologies are used in order to obtain relevant data analysis engines.

The integration of tools and technologies for building IDA processes is a key issue for developing applications that improve the effectiveness of the e-Learning platforms. The EDM community will benefit from the discussions related to the advantages and drawbacks of various options in a practical context with experimental results, by improving the efficiency of building high quality software systems supporting the research efforts.

The first step of developing an IDA application should focus on choosing the right tool or technology that fits the task requirements (e.g., input size, algorithm type, running time efficiency, scalability, etc.). The diversity of the available options is an indication of the necessity for a detailed analysis. From this point of view, the EDM community needs to be aware of success and failure attempts of many practical research efforts in order to provide the possibility of a proper future design choice.

Existing tools and technologies implement in different ways recent advances on techniques from statistical/machine learning, information retrieval and data mining domains in terms of programming language (e.g., Java, C/C++, C#, R, etc.), toolkits (e.g., Weka, Apache Mahout, MLTK, Maple, Matlab, etc.) and implementation details that may have a great impact on running times, scalability, effectiveness or efficiency.

This workshop brings together researchers from academia and industry practitioners with special interest in statistics/machine learning, information retrieval, data mining to (1) discuss current state of the art tools and technologies, (2) identify patterns for proper usage of various options for different tasks, and (3) lay out a vision regarding the modality in which tools and technologies will influence future applications. The organizers hope to obtain common background knowledge for integrating various tools and technologies in future EDM applications.

Vladimir Fomichov, National Research University Higher School of Economics (HSE), Moscow, Russian Federation
Costin Bădică, University of Craiova, Romania
Innar Liiv, Tallinn University of Technology, Estonia
Adrian Graur, University of Suceava, Romania
Vladimir Ivančević, University of Novi Sad, Serbia
Vladimir Cretu, University "Politehnica" of Timisoara, Romania
Mihai Garboveanu, University of Craiova, Romania
Ivan Chorbev, Ss. Cyril and Methodius University Skopje, R. of Macedonia
Ion Iancu, University of Craiova, Romania
Liana Stănescu, University of Craiova, Romania
Marius Brezovan, University of Craiova, Romania
Urszula Markowska-Kaczmar, Wroclaw University of Technology, Poland

SMLIR Workshop Chair
Marian Cristian Mihăescu

Steering and Organizing Committee
Mihai Mocanu
Costel Ionascu
Mirel Cosulschi