



# International

Innovation in Knowledge Based and Intelligent  
Engineering Systems



## INVITED SESSION SUMMARY

### Title of Session:

(DECREASE) DELving into imbalanCed, oveRlapping and multidimEnsionAl data: recent advanceS and challEnges

### Name, Title and Affiliation of Chair:

#### Co-Chairs:

- Małgorzata Bach, PhD, Department of Applied Informatics, Silesian University of Technology, Poland
- Aleksandra Werner, PhD, Department of Applied Informatics, Silesian University of Technology, Poland
- Daniel Kostrzewa, PhD, Department of Applied Informatics, Silesian University of Technology, Poland

### Details of Session (including aim and scope):

The real data on the basis of which scientists try to build decision models and draw conclusions are often not perfect. They can be noisy, incomplete, biased, high dimensional, or, for example, non-representative. All this causes that they require costly and time-consuming processing.

One of the main issues researchers face in this regard is imbalanced data. This problem can be observed in many areas, e.g., in the medical diagnosis of rare diseases, where the number of patients suffering from such diseases is very low in the population, in the detection of fraud in card transactions, where the number of legitimate transactions is much higher than the number of fraudulent ones, and so on.

In most of the mentioned cases identifying rare objects is of crucial importance, unfortunately, the unequal distribution of classes within a dataset causes problems in constructing machine learning models that would be able to correctly classify such (rare) objects.

Another commonly known data problem is the curse of dimensionality, which concerns the exponential growth of necessary experimental data as a function of the dimension of space. In many cases, high dimensionality does not provide additional information but also increases computational complexity and degrades the generalization of the detection model. Therefore, it is often necessary to use techniques involving dimensionality reduction and feature selection.

The session *"Delving into Imbalanced, Overlapping and Multidimensional Data: Recent Advances and Challenges"* aims to enable researchers to exchange knowledge and experiences, as well as to discuss the possibilities of solving a broad spectrum of problems caused by class overlap, imbalance and high-dimensional data.

Therefore, the topics of interest include, but are not limited to:

- Learning from imbalanced and/ or overlapping data
- Learning from high dimensional data
- Dimensionality reduction methods
- Feature selection and feature extraction techniques
- Handling class imbalance in real-world datasets
- Undersampling methods
- Oversampling methods
- Evolutionary computation
- Machine learning
- Deep learning strategies
- Pattern recognition
- Data compression

### Main Contributing Researchers / Research Centres (tentative, if known at this stage):

Silesian University of Technology, Poland  
Kaunas University of Technology, Lithuania,  
University of Economics, Prague, Czech Republic,  
University of Rijeka, Croatia

**Website URL of Call for Papers (if any):**

**Email & Contact Details:**

- Małgorzata Bach, malgorzata.bach@polsl.pl
- Aleksandra Werner, aleksandra.werner@polsl.pl
- Daniel Kostrzewa, daniel.kostrzewa@polsl.pl