- **Title and acronym of the special session** : Incremental Machine Learning (IML)

- **Description of the topic**

  Recently, the size of available datasets is increasing rapidly due in part to the development of internet and sensor networks. In many cases, these databases are constantly evolving (data streams). They are characterized by a changing structure over time, new data arriving continuously. Sometimes, the evolution and the mass of data is so important that it is impossible to store them in a database. This poses several unique problems that make obsolete the applications of standard data analysis. Indeed, these databases are constantly online, growing with the arrival of new data. Thus, efficient algorithms must be able to work with a constant memory footprint, despite the evolution of the stream, as the entire database cannot be retained in memory. Traditionally, most machine learning algorithms are focused on batch learning from a static dataset or from a well-known distribution. However, these batch algorithms take a lot of time to learn a large amount of training data and many batch learning algorithms are not adapted to deal with non-stationary distributions. Only an analysis "on the fly" is possible. These processes are called "data streams analysis" and are the subject of numerous studies in recent years due to the large number of potential applications in many fields. Online incremental algorithms process few examples at a time and allows to extract the knowledge structure from continuous data in real-time. However, this problem became more difficult when we deal with high dimensional data, unbalanced data or outliers. Indeed, the study of data streams is a difficult problem: the computing and storage cost are high and the size of involved datasets is big. In the field of data mining, the main challenges for the study of data streams are the ability to compute a condensed description of the stream properties, but also the detection of change in the stream structure. To be autonomous, the algorithms must have several important characteristics: they must be able to discover the underlying structure of the data without any prior knowledge, i.e. the algorithm should be independent and automatically learn the spatial and temporal structure of the regardless of the type of the data: vector, symbolic, more complex forms (text, graphs) or mixed. The algorithms have to adapt to structural changes in data over time ("concept drift") in real time, without having to relearn everything every time. They must be able to store and reuse knowledge to learn a new data structure similar to that already learned. This property is essential for data stream analysis.
  This special session aims to act as a forum for new ideas and paradigms concerning the field of Automated Incremental Learning. It will solicit theoretical and applicative research papers including but not limited to the following topics:

  - Incremental Supervised Learning
  - Incremental Unsupervised Learning
  - Online Learning
  - Autonomous Learning
  - Concept Drift
  - Model Selection
  - Online Feature Selection
  - Clustering data stream
  - Distributed Clustering
  - Consensus Clustering
  - Incremental Probabilistically Models
  - Active Learning
  - Applications of Incremental Learning

- **Past records**
  The organizers already organized several workshops and special sessions at WCCI 2012, WCCI 2014, IJCNN 2015 on Incremental Machine Learning and Automated Machine Learning, with a very good attendance record.

- **Short bio of the organizers**

Guénaël Cabanes (lead organizer)
LIPN, Paris 13 University, Villetaneuse, France
cabanes@lipn.univ-paris13.fr

Guénaël Cabanes is an associate professor in Computer Science at the University of Paris 13, member of the machine learning research team at the LIPN-CNRS laboratory. He received a M.Sc. in Computer Science in 2007 and a Ph.D. in Computer Science (Unsupervised Learning) at the University of Paris 13, France, in 2010. His primary research interests are in incremental unsupervised learning, autonomous data mining and complex data analysis. He has published approximately 25 papers in refereed journal and conferences proceedings. He is also a member of IEEE, INNS and INNS AML (Autonomous Machine Learning group).

Nicoleta Rogovschi
LIPADE, Paris Descartes University, Paris, France
nicoleta.rogovschi@parisdescartes.fr

Nicoleta Rogovschi received her Master of Computer Science degree from Paris 13 University in 2006 in Machine Learning. She completed her Ph.D. in Computer Science (Probabilistic Unsupervised Learning) in 2009 in the Computer Science Laboratory of Paris 13 University. She is currently an Associate Professor in Computer Science at the Paris Descartes University. She's research is with the Data Mining (GFD) Team. Her research interests include Probabilistic Learning, Unsupervised Learning, Clustering and Co-Clustering methods for different types of data. She is also a member of EGC, AFIA, IEEE, INNS, INNS AML group.

Nistor Grozavu
LIPN, Paris 13 University, Villetaneuse, France
nistor@lipn.univ-paris13.fr

Nistor Grozavu received his Master of Computer Science degree from Marseille II University in 2006 in Fundamental Informatics. He completed his Ph.D. in Computer Science (Unsupervised Learning) in 2009 in the Computer Science Laboratory of Paris 13 University. He is currently an Associate Professor in Computer Science at the Paris 13 University. His research is with the Machine Learning and Application Team from the LIPN Laboratory. His research interests include Unsupervised Learning, Transfer Learning, Dimensionality reduction, Collaborative Learning, Machine Learning by Matrix Factorization and content based information retrieval. He is also a member of IEEE, INNS, INNS AML group (founder and head of the group).

Younès Bennani
LIPN, Paris 13 University, Villetaneuse, France
bennani@lipn.univ-paris13.fr

Younès Bennani received B.S. degree in Mathematics and Computer Science from Rouen University, in 1987. Subsequently, he received the M.Sc. and the Ph.D. degree in Computer Science from The University of Paris 11, Orsay, in 1988 and 1992, respectively, and the accreditation to lead research in Computer Science from the Paris 13 University in 1998. He joined the Computer Science Laboratory of Paris-Nord (LIPN-CNRS) at Paris 13 University in 1993 as Assistant Professor. In 2001, he was appointed to a Full Professor of computer science in the Paris 13 University. His research interests are in theory of Connectionist Learning (Neural Networks), Statistical Pattern Recognition and data mining. His areas of expertise are unsupervised learning, cluster analysis, dimensionality reduction, features selection, features construction, data visualisation, and large-scale data mining. He has published 2 books and approximately 150 papers in refereed conferences proceedings or journals or as contributions in books. He is a senior member of the IEEE.

- **Tentative program committee**

  - Guénaël Cabanes, Paris 13 University, France
  - Nistor Grozavu, Paris 13 University, France
  - Vladimir Nikulin, Vaytka State University, Russia
  - Nicoleta Rogovschi, LIPADE, Paris Descartes University, France
  - Lazhar Labiod, LIPADE, Paris Descartes University, France
  - Asim Roy, Arizona State University, USA
  - Plamen Angelov, Lancaster University, United Kingdom
  - Agostino Gibaldi, University of Geneva, Italy
  - Vincent Lemaire, Orange Labs, France
  - Shiliang Sun, East China Normal University, China
  - JunBin Gao, Charles Sturt University, Bathurst, Australia
  - Younès Bennani, Paris 13 University, Villetaneuse, France
  - Faiciel Chamroukhi, LSIS, Southern University of Toulon, France