Thesis proposal in Computer Science

Title of the thesis: Adaptive system for supporting analytical process design

Key word: data analysis, knowledge extraction, decision support system, meta-analysis, quality measure, similarity measure

Context: Data analysis contains various methods and techniques of mining (for example, classification, factor analysis, pattern mining, etc.). Data analysis can also be viewed as an iterative and/or incremental process of the application of these methods ([1]): the result of a mining technique potentially leading to be replayed or to be continued by another technique.

Even supposing a user is an expert of the data to analyse, it is usually difficult for him to appropriate all possible techniques to obtain the most relevant result. In order to facilitate data analysis, a number of studies have been devoted to the support of analytical process design and, in particular, using meta-analysis approaches ([2], [3], [4], [5], [6], [7], [8]). Meta-analysis consists in discovering similar approaches among a set of analyses. Especially, it implies an ability to search for patterns of analysis in an efficient and optimal way.

Objectives: One key point, particularly difficult, is that these assistants must integrate each step of the analysis process, helping the end user at best. These steps are the following:
- pre-treatment: ensure the good quality of the input data,
- parametrization of the operators: adjust correctly the parameters of the tasks forming an analytical process,
- post-processing: ensuring the reliability of the analysis results.

The post-processing step is still an important challenge in current decision support systems. Indeed, this step is responsible for the understanding and reliability of the results produced in a real context. The thesis will focus on these fields. The aim will be to propose techniques enabling the end user to appropriate the analysis approaches recommended. This appropriation can be seen in two complementary ways:
  - adapting/customizing, at best, recommendations to the context of the user
  - by allowing the exploration of analytical processes, performed by previous users, using a navigation language. It will imply to propose a definition of analytical process as accurate as possible (in particular, its environment: for example, the input data and the expected results).

Funding: Application for State Fellowship

Duration: 36 month starting from October 2017

Laboratory: Institut de Recherche en Informatique de Toulouse (IRIT), SIG team.

Supervisors: Chantal Soulé-Dupuy – Full Professor in Computer Science (IRIT – Université Toulouse 1 Capitole) et Julien Aligon – Associate Professor in Computer Science (IRIT – Université Toulouse 1 Capitole)

Doctoral School: Ecole Doctorale Mathématiques, Informatique et Télécoms de Toulouse (ED MITT).

Candidate Profile: Master degree (or equivalent) in computer science (data analysis and information system). A profile with strong skills in decision support system will be particularly appreciated. It is also required strong skills on at least one programming language. English proficiency.

To apply, please to send as soon as possible, and before May 08, 2017, the following elements:
  - Curriculum Vitae
  - an application letter outlining the motivations
  - transcripts of the Master’s degree
  - contacts and / or letter of recommendation of educational referents

à chantal.soule-dupuy@irit.fr et julien.aligon@irit.fr
References:

Professional perspectives of the PhD student
During the thesis, the PhD student will benefit from training in research, which will open up broader perspectives for his/her integration, both in the academic field (teacher and researcher in computer science) and in the industrial sector (research and development). The PhD student will be sensitized and supported throughout his thesis for the definition of his professional project. In particular, the IRIT team organizes meetings for PhD students on all the possibilities offered, possible ways, constitution of applications, etc. Also, beyond the possibilities of insertion in public or private research, the PhD student will be able to claim engineering positions (data analysis, data extraction, etc.) in all fields and especially in the field of data science.