

# **Séminaire MIDI : Zaineb Chelly Dagdia**

15 Mai 2020, 11:00

## **Titre du séminaire et oratrice**

New Evolutionary Algorithms based on Machine Learning and Data Mining Techniques in Certain and Imprecise Contexts, and Big Data.

Zaineb Chelly Dagdia, Postdoctorante INRIA Nancy

## **Date et lieu**

Vendredi 15 mai 2020, 11h.

Par visioconférence

## **Résumé**

My research aims at developing new evolutionary classification methods in an imprecise framework. A special focus is given to the Dendritic Cell Algorithm (DCA) by addressing and resolving its issues via an investigation of its algorithmic characteristics. New DCA versions were developed leading to more robust DCA binary classifiers based on machine learning tools. These newly developed methods rely, first of all, on solving the DCA non-robust data pre-processing phase by proposing a set of automated modules based on Rough Set Theory and Fuzzy-Rough Set Theory. Secondly, the developed methods aim at investigating the reasons behind the sensitivity of the DCA to the input class data. This was handled by incorporating fuzzy set theory, a fuzzy clustering technique and a database maintenance technique. Lastly, an overall hybrid DCA was proposed taking into account all the mentioned parts. A distributed DCA was indeed implemented to deal with big data. My research also deals with the application of machine learning techniques to deal with big data which are specified by their Volume, Variety, Velocity and Veracity/Imprecision. Based on these specificities, it has become difficult to quickly acquire the most useful information from the huge amount of data at hand. Thus, it is necessary to perform data pre-processing as a first step. In spite of the existence of many techniques for this task, most of the state-of-the-art methods require additional information for thresholding and are neither able to deal with the big data veracity aspect nor with their computational requirements. I work on filling these major research gaps by proposing solutions based on Rough Set Theory for data pre-processing and Randomized Search Heuristics for optimization. My research also deals with real-world applications mainly in epidemiology and security.

<https://sites.google.com/site/zeinebchelly/home>

