Data characteristics and their relation to closed patterns discovery algorithms

Engelbert Mephu Nguifo
LIMOS, Clermont University, Blaise Pascal University and CNRS
Clermont-Ferrand, France
mephu@isima.fr

Abstract. Closed frequent patterns discovery remains a challenge in data mining. During the last decade, different works on data mining algorithms have based their performance evaluation on one dataset characteristic: its density (or on the contrary its sparseness). The incoming of massive datasets in different applications, points out the important goal to design efficient algorithms. The density measurement have shown to be a direction to reach such goal, especially when dealing with formal context of concept lattices. This talk will discuss this notion and describe some metrics defined to characterize dataset density for patterns discovery purpose.

References