

Title: Integrative biological itemset mining in cancer research

Authors: Naulaerts Stefan^{1,2}, Vanden Berghe Wim³, Laukens Kris^{1,2}

¹Department of Mathematics and Computer Science, University of Antwerp, Antwerp, Belgium

²Biomedical informatics research center Antwerpen (biomina), University of Antwerp, Antwerp, Belgium

³Department of Biomedical Sciences, Lab of Protein Chemistry, Proteomics & Epigenetic Signaling (PPES), University of Antwerp, Antwerp, Belgium

In the last decades, a wealth of data has become available and is ready to be analyzed thanks to careful database design. However, the sheer volume makes it impossible to analyze all this information manually and masks potentially interesting patterns. This problem can partially be overcome by using the advances in machine learning and knowledge discovery as supportive tools for life scientists. Innovative frequent itemset search algorithms are capable of uncovering hidden patterns and can be fine-tuned to focus on capturing biological relevance instead of only the database characteristics. However, this requires that several weaknesses of FIM algorithms will be addressed, such as the dense data problem and the lack of biologically relevant quality measures that can be used to fine-tune the algorithms. Fortunately, many different types of biological information have become available and can be combined to redefine the interestingness criterium to the life sciences environment.

In this poster, we present a frequent itemset mining (FIM) framework, powered by information from public and in-house repositories as an assistant platform for integratomics analyses. As such, we build on existing techniques used in pathway and functional enrichment and combine these with biologically relevant modifications of current state-of-the-art data mining techniques. We hereby tackle several of the traditional shortcomings of FIM algorithms, while validating and applying our multi-level approach to currently ongoing cancer research to identify regulatory systems.