ICON-Bio: Integrated Connectedness for a New Representation of Biology

Description

The aim of the project is to develop a comprehensive framework for generalizing network analytics and fusion paradigms of non-negative matrix factorization to medical data. Heterogeneous, interconnected, systems-level omics data are becoming increasingly available and important in precision medicine. We are seeking to better stratify and subtype patients into risk groups, discover new biomarkers for complex and rare diseases, personalize medical treatment based on genomics and exposures of an individual, and repurpose known drugs to different patient groups. Existing methodologies for dealing with these big data are limited and a paradigm shift is needed to achieve quantitatively and qualitatively better results. The project is motivated by the recent success of non-negative matrix tri-factorization (NMTF) based methods for fusion of heterogeneous data in biomedicine. Though these methods have been known for some time, the availability of large datasets, coupled with modern computational power and efficient optimization methods, allowed for creation and efficient training of complex models that can make a qualitative breakthrough. For example, NMTF has recently achieved unprecedented performance on exceptionally hard problems of simultaneously utilizing the wealth of diverse molecular and clinical data in precision medicine. However, research thus far has been limited to special variants of this problem and used only fixed point methods to address these exciting examples of hard non-convex high-dimensional non-linear optimization problems. The ambition of the project is to develop general data fusion methods, from mathematical models to efficient and scalable software implementation, and apply them to the domain of biomedical informatics. The project will lead to a paradigm shift in biomedical and computational understanding of data and diseases that will open up ways to solving some of the major bottlenecks in precision medicine and other domains.

Barcelona Supercomputing Center - Centro Nacional de Supercomputación