Moving MS-based Data Science from raw Spectra towards comprehensive Findings

Today's challenges in MS-based data analysis include determination of the confidence of identification and quantification, integration of data across various omics platforms, and assurance of FAIR storage of predictive models derived from study results. Here, I present the vision of a novel predictive framework wrapped in a comprehensive analysis pipeline to address these issues. My vision includes the development of efficient algorithms for rapid interpretation of raw MS data with statistical confidence, and extends to integrative methods capable of merging datasets from proteomics, lipidomics, and metabolomics for true multiomics analyses. These methods and models providing the study outcome will then be stored in a novel database framework for prediction / classification models. The so-called Bio Classifier is a centralized database designed to store classification and predictive models originating from diverse analytical domains, e.g., biological processes, metabolic dynamics, nutrition, medical health states, and environmental conditions. This setup enables rapid comparisons between newly generated data and existing models, allowing users to quickly identify correlations with other conditions for truly comprehensive findings and study outcomes.