

Student Project

FEATURE ENGINEERING FOR MICROBIOME DATA

Student's background: computer science, biomedical engineering, bioinformatics, biotechnology, mathematics, medicine, or other fields with analytical background with interest in computational biology.

Type of project: BSc/MSc

Project description: This project aims at finding and evaluating quantitative methods to extract relevant features from microbiome data to be used for machine learning or statistical analyses. There are key challenges in this project: first to deal with high dimensionality of the data itself that renders the extraction of novel features computationally demanding, second the issue of compositionality of microbiome data in case of using relative taxa abundances, and the interpretability in taxonomic-based feature extraction versus a non-referenced data mining approach from raw metagenomic data.

Methods used: There are a wide variety of quantification methods to use based on basic statistics (e.g. describe data distribution), to advanced methods based on nonlinear dynamics, game theory, evolution and adaptation-based methods, and pattern formations. In this process, the student will learn about microbiome data and feature engineering techniques. The student will be expected to develop new features or adopt novel features in this area.

About PERSIMUNE: PERSIMUNE (www.persimune.dk) is a multidisciplinary centre of excellence embedded within CHIP and funded by the Danish National Research Foundation. PERSIMUNE works from the hypothesis that across patients with impaired immune function, there is a common pattern of un-discovered risk factors explaining the variation in risk of infectious complications. We aim at understanding the mechanisms explaining the variation in risk using a diverse set of methodologies, including pattern recognition from big data from routine clinical care, studies of host and microbial genetics, imaging, and immunological characterization. These data will be used to develop clinical algorithms aimed at improving the clinical outcomes in patients with various immune dysfunction.

Supervisor: Ramtin Zargari Marandi

Email: ramtin.zargari.marandi@regionh.dk