

Predicting microbiome changes over time and upon perturbation: applications of machine learning to microbiome research and therapeutic strategies

The human body forms different ecological niches for microorganisms. We host a plethora of microbes on each part of our body – different parts of our skin, our mouths; and in our body, especially in the gut. Those different niches are called microbiomes, for example the gut microbiome. In a bit over a decade the scientists described different microbiomes and learned that the microbiome is not only necessary, as it co-evolved with us, diverse and complex, but also important to health. Changes in the composition of microbes may result in an imbalance within the ecosystem of our guts called dysbiosis, which in turn may lead or contribute to disease. We already know a wide array of microbiome-disease associations, from the quite expected ones, such as obesity or gastrointestinal problems (ulcerative colitis or Crohn's disease), to diseases presumed far removed from the gut – type-1 diabetes, allergies, depression, anxiety, neurodegenerative disorders, cancer, and others.

With those significant health implications and vast amounts of data gathered over the years we are now in a position to turn the knowledge into action – learn to predict how the microbiome changes over time and upon perturbation. With this, we are going to learn how to change the microbiome in order to promote health or devise therapeutic strategies. Those strategies may involve probiotic interventions, dietary or lifestyle changes, not necessarily classical small-molecule drugs. This is no easy task as the microbiome is massive – 10s of trillions of microbial cells representing 100s of microbial species, harboring over 100 times more genes than the human genome. On top of that the microbiome is dynamic. Unlike our genome, which is fixed from birth, the microbiome changes from day to day, evolves with our diet, lifestyle and other life events.

Hence, in this project we will employ a cautious step-wise strategy to pave the way for future smart microbiome-oriented therapies. As a first step, we will construct computational models of how the microbiome changes from day-to-day. Analyzing long time-series data from several individuals we are going to construct statistical and machine learning-based models (auto-encoders) to predict those changes. Thanks to those models, we will learn what features (microbes or combinations of microbes) are important for microbiome evolution over time. We will also better learn which specific algorithms and neural network architectures work for this problem. Once we master this task, we will be ready to proceed to a more complex one.

One of the few microbiome-oriented therapeutic strategies are the fecal microbiome transplants (FMTs). Initially, they were used for the treatment of *Clostridium difficile* infections, but now they also find their use in the treatment of ulcerative colitis. Collecting the data from already published studies, we will predict the composition of patient's microbiome after an FMT. This is a substantially more complicated task, so we will use previously designed models for time-series data to give us a head start (transfer learning paradigm).

As a result of this project, we will learn how to predict the human gut microbiome changes over time and in response to FMTs. This will be a first step towards intelligent microbiome-oriented therapeutic strategies which give significant hope for many of the most debilitating diseases of today.