

Personalized gut microbiome analysis for colorectal cancer classification with explainable AI

March 29 2023

Personalized Gut Microbiome Analysis with AI for Classifying Colorectal Cancer



Machine learning models can be used to discover potential colorectal cancer (CRC) biomarkers in the gut microbiome



However, most of these models cannot recognize bacterial species relevant only for smaller patient groups

Explainable AI framework for personalized feature importance in CRC prediction





Clear separation

between healthy

in 2D feature

space

and CRC subjects



Contribution of each input feature



Contribution of each bacteria species to each CRC prediction



Identification of CRC subgroups differing in CRC probability and bacterial abundance



This approach will enable the exploration of personalized microbiome data, uncovering disease subgroups and their biomarkers

Leveraging explainable AI for gut microbiome-based colorectal cancer classification Rynazal *et al.* (2023) | *Genome Biology* | 10.1186/s13059-023-02858-4





Graphical abstract. Credit: Tokyo Tech

The gut microbiome comprises a complex population of different bacterial species that are essential to human health. In recent years, scientists across several fields have found that changes in the gut microbiome can be linked to a wide variety of diseases, notably colorectal cancer (CRC). Multiple studies have revealed that a higher abundance of certain bacteria, such as Fusobacterium nucleatum and Parvimonas micra, is typically associated with CRC progression.

Based on these findings, researchers have developed various artificial intelligence (AI) models to help them analyze which bacterial species are useful as CRC biomarkers. However, most of these models rely on what is known as "global explanations," meaning that they can only consider the entirety of the input data to make predictions. As a result, such models are unable to identify bacterial species that could be relevant CRC biomarkers for smaller, less-representative groups of patients.

Against this backdrop, a research team from Tokyo Institute of Technology (Tokyo Tech), Japan, decided to adopt a different approach capable of addressing this limitation. As outlined in their paper, which was recently published in *Genome Biology*, the team employed an explainable AI framework that provides local, rather than global, explanations for its CRC predictions.

"Local explanation techniques make it possible to discover the most contributing bacteria for each individual CRC patient, enabling us to examine inter-individual differences between subjects within a disease group," explains Associate Professor Takuji Yamada, the main author of



the study.

The team used a framework called "Shapley additive explanations" (SHAP), which originated from a concept in <u>game theory</u> called the Shapley value. Put simply, the Shapley value tells us how a payout should be distributed among the players of a coalition or group. Similarly, in their study, the team used SHAP to calculate the contribution of each <u>bacterial species</u> to each individual CRC prediction.

Using this approach along with data from five CRC datasets, the researchers discovered that projecting the SHAP values into a twodimensional (2D) space allowed them to see a clear separation between healthy and CRC subjects. Clustering this 2D information resulted in four distinct subgroups of CRC subjects, each differing in the CRC probability and the associated bacteria.

In addition, the team found that subjects in the CRC subgroups with the highest CRC probability always had an enriched population of bacteria typically associated with CRC. Most remarkably, the results were consistent across the five datasets, showcasing the wide applicability of this method.

With these promising results, the team anticipates their approach to make solid contributions in the gut microbiome research community.

"Considering the increasing use of machine learning in microbiome–disease association studies, our novel method could be beneficial for a more personalized <u>microbiome</u> data exploration as well as help uncover potential disease subgroups along with their potential associated biomarkers," speculates Dr. Yamada. Further, the technique is also applicable to other diseases with known links to the <u>gut microbiome</u>, such as <u>ulcerative colitis</u>, Chron's disease, and diabetes.



More information: Ryza Rynazal et al, Leveraging explainable AI for gut microbiome-based colorectal cancer classification, *Genome Biology* (2023). DOI: 10.1186/s13059-023-02858-4

Provided by Tokyo Institute of Technology

Citation: Personalized gut microbiome analysis for colorectal cancer classification with explainable AI (2023, March 29) retrieved 26 April 2024 from <u>https://medicalxpress.com/news/2023-03-personalized-gut-microbiome-analysis-colorectal.html</u>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.