Writeup - sbv IMPROVER Metagenomics Diagnosis for IBD Challenge

SUBCHALLENGE: ==X== (==PLEASE, PRECISE IF 1 OR 2==)

Date: ==XXX==

==PLEASE DO NOT ADD ANY CONTACT DETAILS AND NAMES TO ENSURE ANONIMITY OF YOUR SUBMISSION==

==IMPORTANT NOTE==: all program names, versions and parameters used for any of the section below should be reported in this document in support of results interpretation, and transparency and reproducibility.

==Your write-up should contain descriptions for the following sections below.==

TABLE OF CONTENT

Data used for training

Were any additional data used? Yes/No

If yes, please, provide a description and source of data

If the answer varies depending on the 2-class problems, please, provide a separate description for each of them.

For data suggested to be used for training by the challenge organizers, please, specify if additional samples (to the ones indicated in the class labels txt files) have been used or some samples have been excluded. Please, explain why. Were all same samples used for all 2-class problems? If not, please, specify.

Data processing

Did you process data in a specific way, and how?

e.g.

for subchallenge 1, please, describe how you processed the sequencing data: QC, profiling techniques, other

for subchallenge 2, please, provide a description if you have applied a specific transformation to data provided as taxonomy and pathway abundance matrices

Approach to train your classification model

Please describe the approach(es)/method(s). Include the algorithm name, software title and version, and parameter values.

Please, provide a description in case parameter values had to be learned/optimized (default parameters were not used). In case different approach(es)/method(s) were used for the different 2-class problems, please, provide a separate description for each of them.

Please, provide the code or at least pseudo-code to allow reproducibility.

Feature selection

Classification model

Application to the testing dataset

Please, provide a description.

Additional information

Please, feel free to add any information that will be helpful for the interpretation of the results and for transparency and reproducibility.