Writeup - sbv IMPROVER Metagenomics Diagnosis for IBD Challenge

Contents

SUBCHALLENGE: 2
Date: 2020-01-14

Data used for training

Were any additional data used? **NO**

Were all same samples used for all 2-class problems? **YES**

Data processing

Data has not been processed as all predictors had same scale.

Approach to train your classification model

We used the randomForest algorithm (Liaw and Wiener 2002).

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.

References

Liaw, Andy, and Matthew Wiener. 2002. "Classification and Regression by randomForest" 2: 6.