

# ML4MICROBIOME WORK PACKAGE 1

## State-of-the-art evaluation and update

Jaak Truu  
University of Tartu  
Estonia

## Objectives

To continuously evaluate the state-of-the-art ML/statistics methods, and to ensure that every action member is “on the same page” in terms of their robustness and suitability for microbiome research and how well they address the specific challenges in 1.1.1, separately and combined.

Task 1.1: Technology watch

Task 1.2: Evaluation of ML/statistics methods currently used in microbiome research

Task 1.3: Define priority areas for novel ML/statistics applications for microbiome data

## WG1 activities

- A1.1: A 'tech watch' group was formed by WG1 members who will closely monitor recent ML/statistics publications, conference disseminations and software releases.
- A1.2: New methods identified by the 'tech watch' will be added to a list of existing ML/statistics methods. These will be continuously evaluated in terms of functionality (e.g. classification, prediction, causality) and suitability for tackling the challenges in 1.1.1, separately and combined.
- A1.3: Based on the outcome of A1.2, priority areas for novel ML/statistics applications that better address the specific challenges will be defined.

## Main decisions at previous action meetings

- The literature/software watch will concentrate on the application of ML in microbiome studies related to causality and clinical use for **diagnostics, prognostics and therapeutics**.
- The literature/software review will be based (in addition to WG1 teams) on crowdsourcing approach ie action participants can submit relevant publications/software.
- In order to facilitate the process of literature review, the web-based solution was set up using feature classification taxonomy for analyzing the research papers.

## WG1 progress

- The literature and software review is based on the combination of the several approaches
  - data input from targeted search by teams within the workgroup 1
  - STSMs
  - input by action members

## Deliverables

- D1.1: Initial evaluation report on methods available at Action start.
  - <https://docs.google.com/document/d/1-fRCVSNsaskyGLt-QdibezxHxZk0ser0/edit>
- D1.2-5: Updated annual reports on recent methods and their performance relative to state-of-the-art.
- D1.6-7: Two reports outlining priority areas for new ML/statistics methods.



# Rapporteur report

Deliverable	Month deliverable due	Delivery status	Rapporteur Comment
Initial evaluation report and subsequent publication(s) on methods available at Action start	12	Delivered	Therefore the need to develop a whole ranking approach just to select interesting papers relevant to MICROBIOME (and not HUMAN) and ML is not fully justified. This can be done manually and COST Action members can focus on deliverables of the MoU. Papers cited in Unsupervised learning section, while showing application of these methods to human data analysis do not directly translate to microbiome analysis (e.g, classification and prognostic of cancer patients with brain metastasis). The section Available Resources for applying ML to human microbiome studies is relevant, albeit very brief, and of interest to this Action.



## Further activities

- Updated annual report for 2021
  - *Will be delivered in October 2021*
  - *Will be based on revised first annual report*
  - *Compilation of the report will follow the same process as last year*
- List of deliverables – any need for change?
  - *D1.6-7: Two reports outlining priority areas for new ML/statistics methods → One reports outlining priority areas for new ML/statistics methods*
- Additional task
  - *Next action/WG1 publication – will be discussed at next WG1 meeting*

The image shows a Zoom meeting interface with seven participants in a grid view. The participants are: Enrique Carrillo - IMDEA Food (top left), Jaak Truu (top middle), Vladimir Trajković (top right), Mikhail Kolev (middle left), Kanita Karadzovic Hadzibabic (middle middle), Piotr Przymus (middle right), and Inés (bottom middle). A chat window is open on the right side, showing a list of participants and a message from Jasminka Hasic: "Microbiome Support, ELIXIR, FNS-Cloud, or others?". The chat also shows a question from Doménica D'Elia: "Is there a web site of the project?" and a response from Jasminka Hasic: "jasminka.hasic@ssit.edu.ba please reach out". A link is shared by Piotr Przymus: <https://www.hmpdacc.org/hmp/resources/>. The Windows taskbar is visible at the bottom, showing the search bar, taskbar icons, and system tray with the date 07/07/2021 and time 13:04.

Thank you!