



MICROBIOME



WG4 Report

A summary of 2019/2023 achievements and work in progress

08 June 2023, Cork University College

COST Action CA18131 - MC&WGs Meeting, Cork, 7-9 June 2023

WG4 Leader: **Domenica D'Elia**

Vice WG-L: **Aldert Zomer**





MICROBIOME



WG4 Objectives

Disseminate the results of the COST Action in

- peer-reviewed journals
- on the web-portal
- at international conferences
- on social networks (LinkedIn and Twitter)
- through end-user workshops & training schools in microbiome-related ML methodologies recommended by Action members

Tasks

T4.1: Disseminate evaluation, optimisation and standardisation progress

T4.2: Organise training courses (e.g. Summer Schools) and workshops





MICRO

HOME

ABOUT US

NEWS

ACTIVITIES AND EVENTS

PUBLICATIONS

TRAINING MATERIAL

GET INVOLVED



ML4MICROBIOME ACTION OVERVIEW

READ MORE

- Presentations
- Meetings & Workshops
- Training
- STSM Grants
- ITC Grants
- VNS Tools
- Dissemination Grants

COST Action CA18131

Statistical and machine learning techniques in human microbiome studies

READ MORE

[ML4Microbiome website](#)





MICROBIOME



Activities:

A4.1: dissemination - COST progress and outcomes shared **within and **outside** the Action at the Web-portal**

✓ [Publications - ML4 Microbiome](#) (12 and many others in progress including one perspective article)

✓ [Presentations - ML4 Microbiome](#) (International Conferences – more than 15)

CONTACT INFORMATION

Address
Grant holder institution:
(GH Manager: Dr Aoife Tierney)
Biosciences Institute,
University College Cork,
Western road,
Cork, Ireland,
T12 YT20.
aoife.tierney@ucc.ie

- Pfeil J, Siptröth J, Pospisil H, Frohme M, Hufert FT, Moskalenko O, Yateem M, Nechyporenko A. **Classification of Microbiome Data from Type 2 Diabetes Mellitus Individuals with Deep Learning Image Recognition**. *Big Data and Cognitive Computing*. 2023; 7(1):5-1. <https://doi.org/10.3390/bdcc7010051>
- **Frontiers in Genetics volume – Research Topic: Microbiome and Machine Learning (2022) – 10 articles published. Microbiome and Machine Learning | Frontiers Research Topic (frontiersin.org)**
- Moreno-Indias L, Zomer AL, Gómez-Cabrero D, Claesson MJ on behalf of ML4Microbiome (2022) Editorial: **Microbiome and Machine Learning Research Topic**. *Front. Microbiol., Sec. Evolutionary and Genomic Microbiology*. <https://doi.org/10.3389/fmicb.2022.964921>
- Cekikj Miodrag, Jakimovska Ozdemir, Milena Kalajdziski, Slobodan Ozcan, Orhan Sezerman, Osman Uğur (2022) **Understanding the Role of the Microbiome in Cancer Diagnostics and Therapeutics by Creating and Utilizing ML Models**. *Appl. Sci.* 12(9), 4094; <https://doi.org/10.3390/app12094094>
- Vilne Baiba, Ķibilids Juris, Sikсна Inese, Lazda Ilva, Valciņa Olga, Krūmiņa Angelika (2022) **Could Artificial Intelligence/Machine Learning and Inclusion of Diet-Gut Microbiome Interactions Improve Disease Risk Prediction? Case Study: Coronary Artery Disease**. *Frontiers in Microbiology*. 13. doi:10.3389/fmicb.2022.627892
- Ibrahimi E., Elbere, I., Berland, M., & D'Elia, D. (2022) **Report of the ML4Microbiome workshop 2021 – Statistical and Machine Learning Techniques for Microbiome Data Analysis**. *EMBnet Journal*, 27, e1012. doi:<https://doi.org/10.14806/ej.27.0.1012>
- Rosario D, Bidkhorji G, Lee S, Bedarf J, Hildebrand F, Le Chatelier E, Uhlen M, Ehrlich SD, Proctor G, Wüllner U, Mardinoglu A, Shoale S. **Systematic analysis of gut microbiome reveals the role of bacterial folate and homocysteine metabolism in Parkinson's disease**. *Cell Rep* 2021 Mar 2;34(9):108807. doi: 10.1016/j.celrep.2021.108807. PMID: 33657381.

CONTACT INFORMATION

Address
Grant holder institution:
(GH Manager: Dr Aoife Tierney)
Biosciences Institute,
University College Cork,
Western road,
Cork, Ireland,
T12 YT20.
aoife.tierney@ucc.ie

Presentations at External Conferences

2021

- **GOBLET & EMBnet AGM 2021, 11-15 October 2021**
- **ML4Microbiome Symposium "Grand Challenges of Data-Intensive Science in microbiome & metagenome data analysis and training"**, 14 October 2021

Oral Presentations by:

- Domenica D'Elia – The ML4Microbiome COST Action: aims & state-of-the-art report**
- Leo Lahti – Open data science in microbiome research**
- Enrique Carrillo – Machine Learning & Microbiome for Precision Nutrition**
- Isabel Moreno – ML4microbiome for the advance of precision medicine: what Biomedicine expects from us**
- Eleni Papakonstantinou – Microbiome analysis in precision agriculture**

- **ML4Microbiome Workshop "Statistical and Machine Learning Techniques for Microbiome Data Analysis"**, 15 October 2021



Home > Frontiers in Microbiology > Evolutionary and Genomic Mic... > Research Topics > Microbiome and Machine Learni...

Microbiome and Machine Learning

19k Total Downloads 91k Views

Overview Articles 10 Authors 92 Impact

10 Articles

Sort by: Views Type Date

EDITORIAL
Published on 05 Jul 2022

Editorial: Microbiome and Machine Learning

Isabel Moreno-Indias · Aldert L. Zomer · David Gómez-Cabrero · Marcus J. Claesson

ORIGINAL RESEARCH
Published on 21 Jan 2021

DeepT3_4: A Hybrid Deep Neural Network Model for the Distinction Between Bacterial Type III and IV Secreted Effectors

Lezheng Yu · Fengjuan Liu · Yikraz Hou · Tingting Liu · Binbin Tian

TECHNOLOGY AND CODE
Published on 28 Jan 2021

kernInt: A Kernel Framework for Integrating Supervised and Unsupervised Analyses in Spatio-Temporal Metagenomic Datasets

[Microbiome and Machine Learning | Frontiers Research Topic \(frontiersin.org\)](https://frontiersin.org)

Topic Editors:

Isabel Moreno Indias, Marcus Claesson, Aldert Zomer, David Gomez-Cabrero

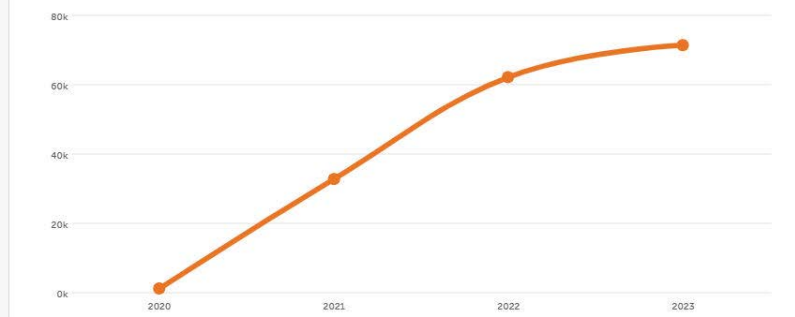
Overview Articles 10 Authors 92 Impact

71265 views

Views Demographics

Since beginning

71265 total views | 68632 views | 19866 downloads | 2633 topic views



Total Views Article Views Article Downloads Topic Views

Full Screen

Since beginning

[Microbiome and Machine Learning | Frontiers Research Topic \(frontiersin.org\)](https://frontiersin.org) - Impact



Microbiome and Machine Learning, Volume II



291
Views

[Participate in this topic](#) →

[Overview](#) [Articles](#) [Authors](#) [Impact](#)

About this Research Topic

Abstract Submission Deadline 30 June 2023
Manuscript Submission Deadline 31 October 2023

[Guidelines](#) >

Due to the success of [Microbiome and Machine Learning](#), which collected research results and perspectives of researchers working in the field of machine learning (ML) applied to the analysis of microbiome data, we are launching the second volume to collate any new findings in the field to further our understanding and encourage the participation of experts worldwide in the discussion.

The success of ML algorithms in the field is substantially due to their capacity to process high-dimensional data and deal with uncertainty and noise. However, to maximize the combinatory potential of these emerging fields (microbiome and ML), researchers have to deal with some aspects that are complex and inherently related to microbiome data. Microbiome data are convoluted, noisy and highly variable, and non-standard analytical methodologies are required to unlock their clinical

Share on   

Participating Journals

Manuscripts can be submitted to this Research Topic via the following journals:

Frontiers in Microbiology
Evolutionary and Genomic Microbiology
Systems Microbiology

Frontiers in Genetics

[Microbiome and Machine Learning | Frontiers Research Topic \(frontiersin.org\) – Volume II \(2023\)](#)

The collection will include:

- Original Research
- Methods
- Reviews
- Technology and Code
- Perspective

Describing machine learning novel tools or approaches for analysing and classifying microbiome data, improving the interpretability of ML application results, or employing Explainable AI

Topic Editors:

Domenica D’Elia, Isabel Moreno Indias, Marcus Claesson, Aldert Zomer, Erik Bongcam-Rudloff, Randi Jacobsen Bertelsen

The table below displays the general rationale of Frontiers' fee policy. Journal-specific APCs are listed on each journal's page, which also contains journal-specific article types. Please note that APCs are subject to periodic revision.


Fees 					
Category	Description	Article type A	Article type B	Article type C	Article type D
1	Journals in fields that have lower research budgets	\$1,260	\$765	\$490	0
2	Journals in fields that support open access	\$2,080	\$955	\$490	0
3	Mature and impactful journals that are well-established and in fields that support open access	\$2,720	\$1,260	\$490	0
4	Sustainable and impactful journals in fields that support open access	\$3,225	\$2,020	\$490	0

Table 1: Frontiers' APCs are structured in four categories, depending on the journal's maturity and the level of research funding and/or open access funding available in various research communities. They also support article types with lower prices or articles that are free of charge. Frontiers' APCs are charged in US\$.

For example, for Original Research articles (article type A), Frontiers' APCs range from **US\$ 3,225** in our most mature journals with well-established open access (OA) support, to **US\$ 1,260** in newly launched journals and/or fields that have lower research budgets available and/or where OA is not yet well-supported (e.g. humanities and social sciences fields). This means that mature journals in well-funded fields have higher APCs and partially subsidize journals with lower APCs that are younger or cover communities with less funding. This allows all research communities to benefit from open access.

Fee discounts

Our fee support program ensures that all articles that pass peer review can benefit from open access – regardless of the author's field or funding situation.

Authors and institutions with insufficient funding will be eligible for discounts to their APC. This includes authors in countries classified by the [World Bank](#) as low and low-middle income economies.

To apply for fee support, please complete our [fee support application form](#).

Your request for a fee support will in no way influence whether your submission is accepted or rejected for publication.

**ML4Microbiome deadline for article submission:
15-20 June 2023**



MICROBIOME

✓ YouTube Playlists: **97 subscribers**

[\(560\) ML4Microbiome Workshop 2021 – YouTube](#)

[\(560\) ML4Microbiome Symposium 2021 – YouTube](#)

✓ Accessible also from the Action website

[Training Material - ML4 Microbiome](#)

ML4Microbiome
@ml4microbiome213 · 97 subscribers
COST Action CA16131 - Statistical and machine learning techniques in human microbiome studies.

ML4Microbiome - Training in microbiome analysis / Leo Lahti
24 views · 1 year ago

EMBnet
GOBLET & EMBnet AGM 2021 / Day 2 Leo Lahti: ML4Microbiome - Training in microbiome analysis 12 October 2020 14:10 CEST ...

Training in Microbiome Analytics | Training in Microbiome Analysis | Elements of Microbiome Data... 7 moments

Welcome, Brief Description of ML4Microbiome & Introduction to the Workshop - Domenica D'Elia
150 views · 1 year ago

ML4Microbiome
ML4Microbiome Workshop 2021 15 October 2021 12:00 CEST <https://www.ml4microbiome.eu/>

COST Action ML4Microbiome - an opportunity for collaboration on training: Domenica D'Elia
18 views · 2 years ago

GOBLET

Statistical analysis of the microbiome data with R - Eliana Ibrahim
1.9K views · 1 year ago

ML4Microbiome
ML4Microbiome Workshop 2021 - 15 October 2021.

Visualization of microbiome data | Statistical hypothesis testing | Sample size and power analysis... 12 moments

Overview of the microbiome data - Ilze Elbere
205 views · 1 year ago

ML4Microbiome
ML4Microbiome Workshop 2021 15 October 2021 12:15 CEST <https://www.ml4microbiome.eu/>

Outline | The Microbiome | What is the aim? | Study design | Design - case/control | The perfect... 22 moments



MICROBIOME

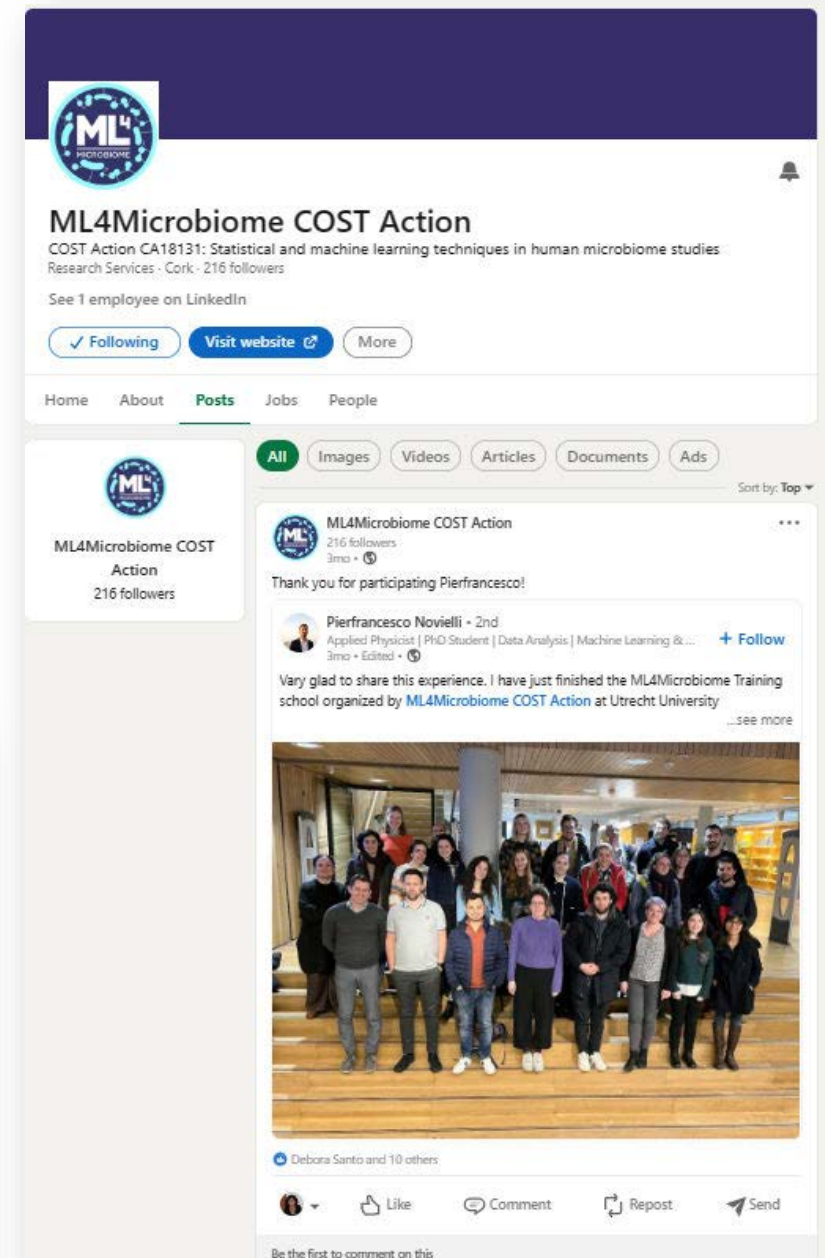
Communication

✓ [LinkedIn ML4Microbiome Channel](#): **216 Followers**

- Visitor highlights June 2022-23: **357**
- Page views: **139**
- Unique visitors: **32**

✓ [Twitter ML4Microbiome Channel](#):

546 Following; 366 Followers



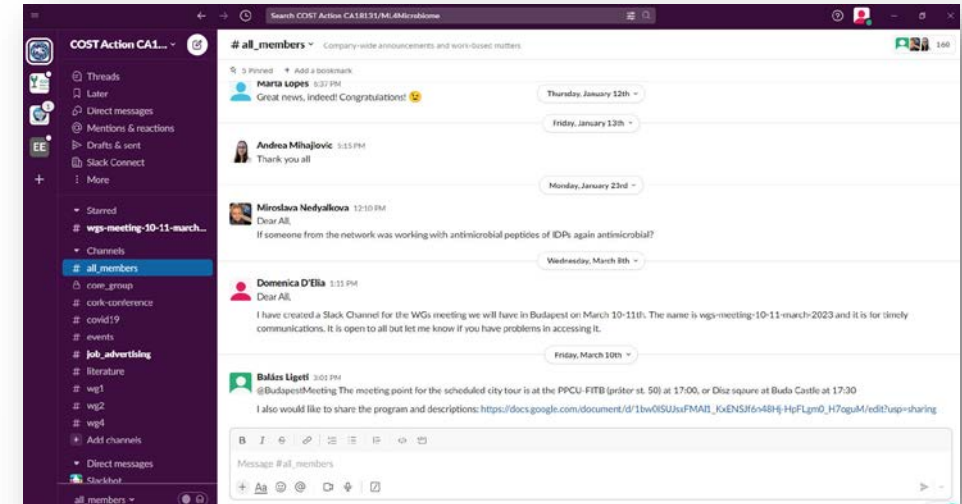
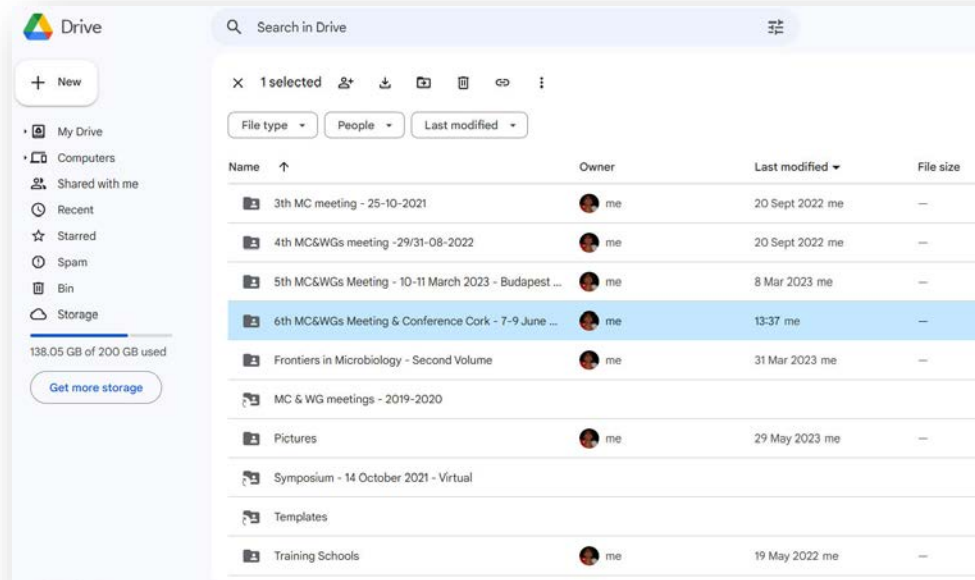


MICROBIOME



Internal Communication & Document Sharing

- ✓ Slack Channel: 160 members
- ✓ GoogleDrive
- ✓ Gmail and COST mailing service





MICROBIOME



A4.2: Design, organisation and implementation of regular training courses and workshops

Milestones:

M4.1: Initial workshop (Y1Q4), training courses (Y2Q3/Y3Q3) and YouTube tutorials (Y4Q3)

Achievements:

✓ **Training Schools** (2019; 2021; 2022; 2023) – Programme and training material available on the **Action website, YouTube and GoogleDrive ML4Microbime accounts**

[\(Activities and Events - ML4 Microbiome\)](#)

✓ **Workshops** ([August 2022, Turku](#) – [May 2022, Tirana](#) – [October 2021, Bari, virtual](#))

✓ **Symposium** ([\(October 2021, Bari, virtual\)](#))

Major Deliverables:

D4.1-3: Reports on the workshops & training courses of best practices in ML/Stats for microbiome studies

D4.4: Educational material shared: tutorials on paper and video (including a YouTube channel)



MICROBIOME



WG4 Report: March-June 2023

- Periodic virtual meetings (every 2-3 weeks)
- Website, LinkedIn and Twitter accounts - timely updated
- [Frontiers in Microbiology: The Second Volume launched](#)
- [ML4Microbiome workshop in BITS 2023 \(21-06-2023\)](#)
- Perspective paper drafting ongoing....to be continued and hopefully concluded on June 9th in Cork (WG4 writing meeting)



MICROBIOME



BITS Bioinformatics Italian Society



Welcome to BITS 2023

19th Annual Meeting of the Bioinformatics Italian Society
June 21-23, 2023, Bari, Italy

The Bioinformatics Italian Society is glad to invite you to its 19th Annual Conference.

The conference highlights keynote talks by excellent scientists in bioinformatics and its applications, presentations of state-of-the-art research in computational biology, and poster sessions on the latest research progress.

The 2023 edition will be held in **Bari 21-23 June**. The event will take place at the [Hotel Excelsior](#) Bari, Via Giulio Petroni, 15, 70124 Bari (IT).



[BITS Meeting 2023](#) ∴ [BITS Bioinformatics Italian Society](#)

Workshop

Advanced machine-learning approaches for the analysis of microbiome data
Workshop

The analysis of the human microbiome has recently attracted the attention of several research communities, due to its potential diagnostics, prognostics, and therapeutics role for several diseases, including diabetes, liver cirrhosis, some types of cancer (e.g., colorectal cancer) as well as for disorders like the autism spectrum disorder. The adoption of statistical and Machine Learning approaches appears very promising to elucidate existing (or identify novel) relationships between microbiome conditions and diseases or to build descriptive and predictive models that can be adopted to improve existing therapeutic procedures. The workshop "Advanced machine learning approaches for the analysis of microbiome data" will focus on advanced machine learning approaches and their (potential or actual) application to microbiome data, including semi-supervised learning methods, multi-view learning methods, and transfer learning approaches.

Workshop Chairs

- **Michelangelo Ceci**, University of Bari, Italy
- **Gianvito Pio**, University of Bari, Italy
- **Domenica D'Elia**, CNR-Institute for biomedical Technologies, Bari, Italy



WELCOME TO BITS 2023

CALLS

WORKSHOP



Andrea Simeon, Mathematician and Data Science researcher at BioSense Institute, Novi Sad, Serbia. Focused on applying Machine Learning and Deep Learning techniques in microbiome studies and exploring different preprocessing pipelines for analysing amplicon and shotgun sequence data. Thanks to the research work carried out during her ML4Microbiome COST Action STSM, Andrea won from the Faculty of Sciences, University of Novi Sad, Serbia, the Aleksandar 'Saša' Popovic Award For Best Student Paper – 2022.

Talk title Multi-class boosting with adversarial multi-arm bandits on incomplete microbiome views

Abstract Microbiome has been massively associated with different diseases and disorders. To identify individual microorganisms and their abundances across samples, different sampling, sequencing and preprocessing techniques could be considered. This leads to different input feature sets (views) to learn predictive models through machine learning (ML) approaches. ML models aid in finding the associations between microbiome and disease. Standard (single view) ML models are not capable of dealing with multiple views: at once, and thus, they were upgraded to fit multi-view datasets (e.g. Adaboost and Multi-view Adaboost). Moreover, microbiome data comes from various sources, and incompleteness is often inevitable. Existing classifiers, even multi-view, cannot be directly used because they cannot work with incomplete views and in multi-class settings. To our knowledge, there is no multi-view boosting algorithm for multi-class classification with incomplete views. The proposed algorithm is the extension of an existing multi-view boosting algorithm based on multi-arm bandits, now able to work in multi-class setting and with incomplete views (views with missing sample representation). At each iteration, it proclaims one view as the winning using adversarial multi-arm bandits and uses its predictive information to update the final model weights and prediction in a boosting process. Three data sets were created from several microbiome studies and used to examine the performance of the proposed algorithm. One of the experiments showed a 7% increase in F1 score compared to a single view classifier, while the other showed 54%. The application domain is not restricted to microbiome data. Further work will involve examinations in other domains.

[BITS Meeting 2023: Workshop](#)

Programme - 21 June 2023

- [09:40-10:00] **Welcome and workshop introduction (Michelangelo Ceci, Domenica D'Elia)**
- [10:00-10:40] **Silvio Tosatto**, DOME: recommendations for supervised machine learning validation in biology
- [10:40-11:10] Coffee break
- [11:10-11:40] **Andrea Simeon**, Multi-class boosting with adversarial multi-arm bandits on incomplete microbiome views
- [11:40-12:00] **Adriano Zaghi**, Machine learning models for the detection of antimicrobial resistance using synthetic data
- [12:00-12:20] **Merixell Pujolassos**, coda4microbiome: compositional data analysis for microbiome crosssectional and longitu
- [11:50-12:10] **Donato Romano**, Explainable Artificial Intelligence (XAI) for Microbiome Data Analysis in Autistic Spectrum Disorder



Perspective paper: [PAPER-DRAFT-revised-DD-05-04-2023 - Google Docs](#)

Frontiers in Microbiology, second volume edition:

PERSPECTIVE Paper

Perspective articles present a viewpoint on a specific area of investigation.

They should provide the following:

- 1) Discuss current advances and future directions
- 2) Clear presentation of the authors' perspective
- 3) Accurate presentation and citations of other authors' work
- 4) May include original data as well as personal insights and opinions

Perspective articles are peer-reviewed, have abstract length 250 words, a maximum word count of 3,000 and may contain no more than 2 Figures/Tables.

Authors are required to pay a fee (B-type article) to publish a Perspective article.

Perspective articles should have the following format:

- 1) Abstract
- 2) Introduction
- 3) Subsections relevant for the subject
- 4) Discussion



Perspective paper: [PAPER-DRAFT-revised-DD-05-04-2023 - Google Docs](#)

Titles proposed:

- 1. Bridging the gap between microbiome data and machine learning: Insights from the ML4Microbiome COST Action**
- 2. Bridging the Microbiome-Machine Learning Divide: Insights from the ML4Microbiome COST Action**
- 3. Facing the challenges encountered by Machine Learning methods on microbiome data: Insights from the ML4Microbiome Cost Action**
- 4. On the challenges encountered by Machine Learning methods on microbiome data: Insights from the ML4Microbiome Cost Action**
- 5. Challenges encountered by Machine Learning methods on microbiome data: Insights from the ML4Microbiome Cost Action**
- 6. Machine Learning and microbiome data: Insights to face the challenges from the ML4Microbiome Cost Action**
- 7. Machine Learning Meets Microbiome: Unpacking the ML4Microbiome COST Action**
- 8. The Intersection of Microbiome and Machine Learning: Key Takeaways from the ML4Microbiome COST Action**
- 9. Uniting Microbiome Data and Machine Learning: Lessons Learned from the ML4Microbiome COST Action**
- 10. Insights from the ML4Microbiome COST Action: Navigating the Convergence of Microbiome and Machine Learning**
- 11. The ML4Microbiome COST Action: A Bridge between Microbiome Data and Machine Learning**

Abstract

The rapid development of machine learning (ML) techniques has opened up the data-dense field of microbiome research for novel therapeutic, diagnostic, and prognostic applications targeting a wide range of disorders, which could substantially improve healthcare practices. However, several challenges must be addressed to fully exploit the benefits of ML in this field. In particular, there is a need to establish "gold standard" protocols for conducting ML analysis experiments and improve interactions between microbiome researchers and ML experts. The Statistical and Machine Learning Techniques in Human Microbiome Studies (ML4Microbiome) COST Action is a European network established in 2019 to promote collaboration between discovery-oriented microbiome researchers and data-driven ML experts, aiming to optimise and standardising ML approaches to the analysis of microbiome datasets. This perspective paper presents the key achievements of ML4Microbiome in optimising, standardising and disseminating the use of ML on human microbiome datasets. These achievements include identifying predictive and discriminatory 'omics' features, improving repeatability and comparability, developing automation procedures, and defining priority areas for the novel development of ML/statistics methods targeting the microbiome. The insights gained from ML4Microbiome will help to maximize the potential of ML in microbiome research and pave the way for new and improved healthcare practices.



Perspective paper: [PAPER-DRAFT-revised-DD-05-04-2023 - Google Docs](#)

Presently completed:

1. Abstract
2. Introduction
3. Subsection: The ML4Microbiome Action Plan and Challenges

Discussion

The discussion should be articulated as follow:

1. Key findings in clear and concise language
2. Place our results within the context of previous studies
3. Discuss potential future research
4. “take-away” statements for the readers

Presently:

1. The current state of ML applications for microbiome data analysis (WG1)
2. Benchmark datasets and online repositories
3. Data preprocessing, transformation and normalization: challenges and solutions (still incomplete)
4. Dissemination, communication and training (not sure it is the case of inclusion unless of training)
5. Define priority areas for the novel development of ML/stats applications targeting the microbiome (still incomplete)
6. New emerging areas: Ethical considerations (contributor Ramona Suharoschi)
7. Intellectual property (IP) rights (contributor Ramona Suharoschi)
8. Explainable AI (contributors Ciprian-Octavian, Marcus Frome, Alina, Elena-Simona)



MICROBIOME



Dissemination & Communication Rules

- **Publications** – any article where at least two members are involved can be published with acknowledgements to ML4Microbiome, and communication has to be done to WG4 (Rules at: [Visual identity – COST](#))
- Production of reports and their publication, hopefully in journals or web repositories that provide a DOI



MICROBIOME



THANK YOU SO MUCH

for your attention!

Question?