Applied Machine Learning for human bacteria MALDI-TOF Mass Spectrometry: a systematic review

Lucía Schmidt Santiago[1,*], Alejandro Guerrero-López[1,2], Carlos Sevilla-Salcedo[1], David Rodríguez Temporal[3], Belén Rodríguez-Sánchez[3], and Vanessa Gómez-Verdejo[1,2]

[1]Department of Signal Processing and Communications, Universidad Carlos III de Madrid, Leganés, 28911 Spain

[2]Signal Processing Group, Gregorio Marañón Health Research Institute, Madrid, 28040 Spain

[3]Department of Clinical Microbiology and Infectious Diseases, Hospital General Universitario Gregorio Marañón. Gregorio Marañón Health Research Institute, Madrid, 28040 Spain

[*]Corresponding author. email: lschmidt@pa.uc3m.es

***Abstract*—Bacterial identification, antimicrobial resistance prediction, and strain typification are critical tasks in clinical microbiology, essential for guiding patient treatment and controlling the spread of infectious diseases. While machine learning (ML) and deep learning (DL) have shown immense promise in enhancing MALDI-TOF mass spectrometry applications for these tasks, a comprehensive review from a technical perspective remains lacking. To address this gap, we systematically reviewed 93 studies published between 2004 and 2024, focusing on key ML/DL aspects such as data size and balance, preprocessing pipelines, peak selection methods, algorithms, evaluation techniques, and the availability of open-source data and code. Our analysis highlights the predominant use of classical ML models like random forests and support vector machines, alongside emerging interest in DL approaches for handling complex, high-dimensional data. Despite significant progress, challenges such as inconsistent preprocessing workflows, reliance on black-box models, limited external validation, and insufficient open-source resources persist, hindering transparency, reproducibility, and broader adoption. By addressing these critical gaps, this review provides actionable insights to bridge microbiological and technical perspectives, paving the way for more robust, scalable, and interpretable solutions in diagnostic microbiology.**

***Index Terms*—MALDI-TOF MS, Machine Learning, Deep Learning, Antimicrobial resistance, Antimicrobial susceptibility testing, Bacterial identification, Bacterial typing**

I. Introduction

Microbial identification and characterization are critical processes in clinical microbiology, particularly for diagnosing infections and guiding appropriate treatments. Accurate species identification enables clinicians to select the most effective antibiotic therapy, minimizing the risk of treatment failure and the spread of antimicrobial resistance. Moreover, rapid identification is essential for infection control, allowing timely implementation of isolation protocols and outbreak prevention measures in healthcare settings.

The advent of Matrix-Assisted Laser Desorption/Ionization-Time of Flight Mass Spectrometry (MALDI-TOF MS) in the late 1980s [1] has revolutionized microbiology, becoming a cornerstone technology for bacterial and fungal identification due to its speed, cost-effectiveness, and reliability [2], [3]. Routinely used in diagnostic laboratories, MALDI-TOF MS converts complex mass spectral data into actionable insights, enabling faster and more accurate diagnostics compared to traditional culture-based methods such as colony morphology on agar plates, which rely on visual growth patterns [4]. By analyzing protein mass spectra, MALDI-TOF MS provides species-level identification within 5 minutes, significantly reducing diagnostic turnaround times. Its ease of use, low cost per sample (€0.5 per sample) [5], and high-throughput capabilities have made it an essential tool in clinical microbiology laboratories around the world.

MALDI-TOF MS involves three fundamental stages: ionization, mass analysis, and detection. A chemical matrix is mixed with the sample to facilitate energy absorption during laser irradiation, resulting in desorption and ionization of sample molecules. These ionized molecules are accelerated through an electric field within a Time-of-Flight (TOF) mass analyzer, where ions are separated based on their mass-to-charge (m/z) ratio. Lighter ions reach the detector faster than heavier ones, producing a time-resolved signal that is converted into a high-dimensional mass spectrum [6], [7]. These spectra, represented as peaks at specific m/z values, serve as unique molecular fingerprints for microbial identification, with each spectrum often containing over 20,000 data points [6], [8].

In addition to identification, effective treatment depends on understanding antimicrobial resistance (AMR) profiles. Antibiotics remain the primary method for treating bacterial infections, however, the rise of AMR presents significant challenges for infection treatment. Resistant strains frequently render standard therapies ineffective, posing a severe global public health threat [9]. In such cases, clinicians rely on antibiotic susceptibility testing (AST) to guide treatment decisions. Methods like disk diffusion, broth microdilution, and automated systems assess bacterial growth in the presence of various antibiotics to identify resistance profiles [10]. How-

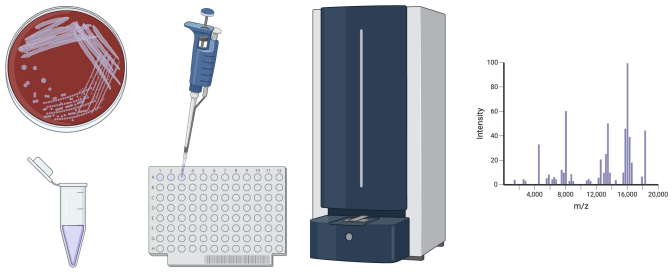

Fig. 1: Worflow of MALDI-TOF MS acquisition.

ever, these methods are resource-intensive and often require 24 to 72 hours for results [11], which can impede timely treatment, exacerbating the burden of nosocomial outbreaks and complicating infection control efforts.

Beyond AMR, bacterial strain typification (TYP), including phenotyping and genotyping, is essential for managing diseases and monitoring epidemiological trends. Phenotyping methods, such as serotyping (ST), rely on observable bacterial traits like surface antigens to classify strains. In contrast, genotyping (GT) focuses on genetic differences, encompassing non-sequencing methods like pulsed-field gel electrophoresis (PFGE), Polymerase Chain Reaction (PCR), as well as sequencing approaches like multilocus sequence typing (MLST), and whole-genome sequencing (WGS) [12], [13]. Ribotyping (RT), a specific GT approach, analyzes variations in ribosomal RNA gene regions and can be performed using non-sequencing methods like PCR-RT or sequencing-based approaches such as WGS-RT [12] useful for identifying hypervirulent strains, associated with nosocomial outbreaks and high mortality rates [14]. While the aforementioned techniques are widely used, they remain cost-intensive (€100-per-sample cost of WGS) and technically demanding.

In recent years, advancements in machine learning (ML) and deep learning (DL) have demonstrated great potential in automatising these challenges. These technologies enable automated bacterial species identification (ID) [15]–[19] and AMR profile prediction [20]–[24] with high accuracy, reducing resource burden. Algorithms such as Support Vector Machines (SVM), Random Forests (RF), and Convolutional Neural Networks (CNN) uncover complex patterns in spectral data that are inaccessible to conventional methods [25]. Such advancements address the pressing need for precise and rapid diagnostic tools in clinical settings [26], [27]. In terms of strain characterization, hence TYP, MALDI-TOF MS, combined with advanced analytical tools, provides a promising alternative for efficient strain characterization and outbreak management [28]–[32].

Despite advancements in this area, several challenges remain. The field requires large, high-quality datasets, standardized pre-processing methods, and strategies to address data imbalance, particularly in AMR-related tasks. Emerging technologies such as generative artificial intelligence (GenAI) and advanced DL models present opportunities to mitigate these issues through improved data augmentation (DA), enhanced model interpretability, and expanded applications of MALDI-TOF MS in clinical microbiology [26], [27]. Furthermore, the development of open-access spectral repositories like the Database of Resistance Information on Antimicrobials and MALDI-TOF Mass Spectra (DRIAMS) [33] and the Robert Koch Institute (RKI) [34] databases underscores the importance of collaborative data sharing to accelerate innovation in the field [22].

In this review, we explore the transformative impact of MALDI-TOF MS on microbial diagnostics and its integration with AI-based methods. We examine its applications in ID, AMR profiling, and strain TYP while addressing current limitations and future directions. By leveraging the synergy between MALDI-TOF MS and cutting-edge AI technologies, we envision a future where diagnostic microbiology is faster, more precise, and universally accessible.

A. Motivation

The integration of ML with MALDI-TOF MS has transformed clinical microbiology, particularly in applications such as microbial ID, antimicrobial AMR prediction, and strain TYP. However, despite significant progress, systematic evaluations of these applications often lack a technical ML perspective.

In 2020, *Weis et al.* [11] published a review in *Clinical Microbiology and Infection* on the application of ML to MALDI-TOF MS in clinical microbiology, summarizing studies published up to that year. Their review encompassed 36 studies, providing an overview of ML-based microbial characterization and AMR prediction, focusing on tasks, bacterial species, and general model descriptions from a microbiological perspective. In 2024, *López-Cortés et al.* [35] specifically analyzed ML applications to MALDI-TOF MS for AMR prediction, covering studies published between 2016 and 2024. More recently, *Pennisi et al.* [36] reviewed AI-based methods for AMR, discussing the predictive performance of ML and DL models and their potential integration into clinical workflows.

While these reviews highlight the strong performance of AI-based models, their primary focus is directed toward the microbiology community. Despite their valuable contributions, they fall short in addressing key ML-specific aspects. Critical considerations such as exploratory data analysis (EDA), preprocessing pipelines, the rationale for model selection, and evaluation protocols are often overlooked. Moreover, the reliance on black-box tools, insufficient validation practices, and the lack of open-source data and code limit their utility for engineering and ML researchers. These shortcomings ultimately hinder the transparency, reproducibility, and trustworthiness of their findings. Nevertheless, the increasing number of publications in this field underscores the growing relevance and importance of these tasks, driving further research and innovation.

This paper fills a crucial gap in the literature by systematically examining ML applications to MALDI-TOF MS from a ML-centric perspective. By revisiting the 36 studies reviewed by *Weis et al.* and expanding the scope to include publications from 2020 to 2024, we aim to provide the ML community with a comprehensive and critical assessment. Of the 40 papers analyzed by *López-Cortés et al.*, 30 are included in our review

for ML-specific evaluation. Specifically, we evaluate the state-of-the-art across key dimensions: EDA (e.g., handling of data imbalance, replicates, and sample acquisition), pre-processing pipelines and their standardization, model decision making based on task type (AMR, TYP, or ID), evaluation practices (e.g., inner cross-validation, outer cross-validation, and out-of-distribution validation), public availability of data and code, reproducibility, and applicability to real-world workflows. In addition, we examine evaluation metrics to propose best practices for future ML-driven clinical microbiology studies. This systematic review serves as a roadmap to improve transparency, rigor, and applicability in this interdisciplinary field.

B. Objective

The primary objective of this review is to systematically analyse ML applications for ID, TYP and AMR detection using MALDI-TOF MS until 2024. While Weis et al. [11] conducted a comprehensive literature search, their analysis lacked a strong ML-focused perspective. Therefore, we re-evaluate the included studies from a ML standpoint to provide deeper insights into the methodologies used. Additionally, given the significant advancements in AI-driven approaches since their review, we have expanded our search criteria to incorporate a broader range of ML- and DL-based methods published through December 2024. Specifically, this review seeks to:

- Critically analyze the most commonly used ML techniques and their performance across diverse tasks.
- Investigate the scope of tasks tackled, while assessing trends in data size, instrumentation, and the range of bacteria and antibiotics studied.
- Examine pre-processing techniques, feature selection strategies, and how data imbalance has been addressed in these studies while highlighting the lack of homogenous pre-processing pipelines.
- Underscore and critique the widespread reliance on black-box software solutions, discussing their limitations in transparency, interpretability, and reproducibility.
- Assess the quality of evaluation metrics, cross-validation methods, and external validation of published results, highlighting inadequate methodological procedures.
- Evaluate the accessibility of datasets and software tools, emphasizing issues of open-source data and proposing strategies to foster a culture of transparency and reproducibility in the field.
- Clearly state the current limitations of ML and DL applications in MALDI-TOF MS and propose actionable future pathways for advancing the field toward clinical integration.

By addressing these objectives, this review aims to provide a critical and constructive perspective on the use of ML in MALDI-TOF MS-based microbial diagnostics. It underscores the need for standardized practices, open-source resources, and transparent methodologies to ensure that ML can be reliably and effectively integrated into clinical workflows while offering a roadmap for future research.

II. Methods

A. Search strategy

A comprehensive literature search was conducted using *PubMed/Medline*, *Scopus*, and *Web of Science*, covering the period up to December 31st, 2024. Following the approach of *Weis et al.* [11], the search aimed to identify studies that applied ML techniques to the analysis of MALDI-TOF MS data. Given the increasing prominence of DL methodologies, the search criteria were expanded to incorporate state-of-the-art DL approaches, ensuring a comprehensive update to current knowledge. The search included specific terms [1] related to ML, DL, and MALDI-TOF MS, applied to titles, abstracts, and keywords where available.

B. Study selection

The study selection process was guided by predefined inclusion criteria, as outlined in the search strategy, and further refined by applying exclusion criteria to ensure the relevance and quality of the studies. The following exclusion criteria were applied:

- **Duplicate records:** Duplicates across databases (*PubMed/Medline*, *Scopus*, and *Web of Science*) were removed during the initial screening phase.
- **Non-full text and irrelevant types:** Records such as review articles, letters, conference proceedings, or those unavailable as full-text articles in English were excluded.
- **Off-topic studies:** Studies that were unrelated to MALDI-TOF MS or did not employ AI-based methods were excluded. This included research focused on genetics, proteomics, cancer, or other unrelated fields.
- **Non-human bacteria:** Articles focusing on non-human bacteria, such as those related to foodborne pathogens, animal studies, viruses, mosquitoes, or fungi, were removed.
- **Off-task focus:** Studies that did not directly address ID, AMR, or TYP were excluded. Examples include studies assessing risk factors, mortality, and morbidity prediction or presenting case studies.

This multi-step exclusion process ensured that the selected studies aligned with the objectives of this review. The PRISMA flow diagram (Figure 2) provides a summary of the screening process and the final set of included studies. A total of 1,888 documents were retrieved from the *PubMed/Medline*, *Scopus*, and *Web of Science* databases. After incorporating 36 studies from *Weis et al.* [11] and removing 871 duplicates, 1,053 unique articles remained for screening. Of these, 128 records were excluded for being non-full-text or irrelevant document types, 378 were deemed off-topic, and 347 focused on non-human bacteria. Additionally, 107 studies were excluded for being off-task. After applying all exclusion criteria, 93 studies published between 2004 and 2024 were included in

[1]The search query included: 'machine learning' OR 'classification' OR 'support vector machine' OR 'random forest' OR 'logistic regression' OR 'neural network' OR 'deep learning' OR 'transformer' OR 'multi-layer perceptron' OR 'embedding' OR 'clustering' OR 'unsupervised' OR 'self-supervised' OR 'semi-supervised' OR 'selfsupervised' OR 'semisupervised' OR 'bayesian' OR 'kernel' AND 'maldi-tof'.

this review. Notably, from the 36 studies reviewed by *Weis et al.* [11], only 25 papers met the updated selection criteria.

All studies are summarized in Table II, which provides an overview of the included works, detailing the bacterial species analyzed, sample sizes, algorithms employed, pre-processing strategies, tasks addressed, and antibiotics tested where applicable. The table highlights the diversity of approaches and focuses across the included studies, offering insights into the evolution of ML and MALDI-TOF MS-based microbial diagnostics over the specified timeframe.

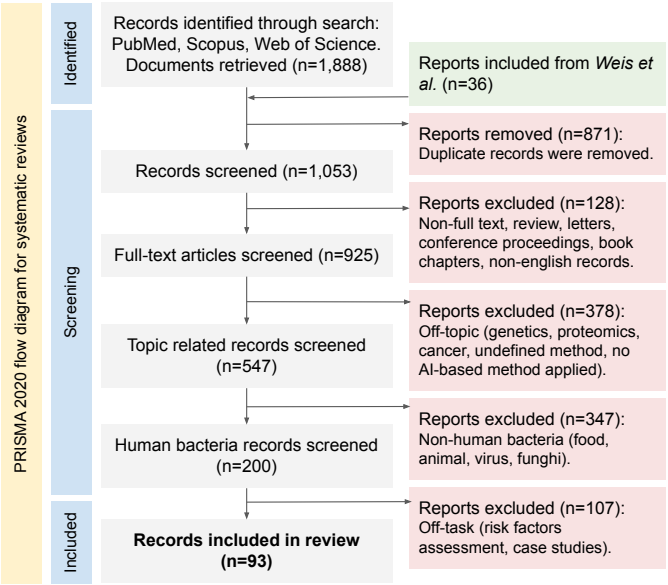

Fig. 2: PRISMA flowchart illustrating the complete literature search pipeline.

C. Quality assessment

In this review, every study underwent a quality assessment based on a predefined set of criteria. This is intended to ensure that the studies are examined in terms of their robustness and clinical relevance. The evaluation focused on key elements such as:

- **ML methodology:** Clear description of the algorithms used, including model architectures, hyperparameters, and training procedures.
- **Cross-validation and external evaluation:** Adequate validation strategies to ensure robustness and generalization, such as k-fold inner cross-validation or the use of external datasets for external or Out-Of-Distribution (OOD) evaluation.
- **Data transparency:** Availability of datasets and/or code to enable reproducibility.
- **Clinical relevance:** The practical application of the findings, particularly in identifying and typing bacterial species and predicting antimicrobial resistance in clinical settings.

D. Data extraction

The following information was extracted from each study to ensure a comprehensive analysis of methodologies and results: (i) **Publication details** included the first author's name, year of publication, and article title; (ii) **Study objectives** were categorized based on whether the study focused on bacterial species ID, AMR detection, or strain TYP; (iii) **Bacterial scope** captured the genera and species investigated in each study; while (iv) **Antibiotics tested** documented the antibiotics analyzed for AMR profiling, where applicable; (v) **Sample size** recorded the number of bacterial isolates or mass spectra used in the analysis.

For computational methodologies, (vi) **ML algorithms** employed in each study were documented, including details of the models, architectures, and specific techniques; (vii) **Pre-processing pipelines** described the steps applied to raw MALDI-TOF MS data, such as normalization, feature extraction, and dimensionality reduction; (viii) **Analysis platforms and tools** identified the programming languages used for implementing the ML models; and (ix) **Performance metrics** included key measures used to evaluate model effectiveness.

To ensure robustness and generalizability, (x) **Validation strategies** were assessed, including internal cross-validation methods (e.g., k-fold) aimed at hyperparameter tuning and OOD evaluation using independent datasets. Finally, (xi) **open-source practices** evaluated the availability of datasets and code to support transparency and reproducibility.

Where available, supplementary materials were reviewed to gather additional information on experimental design, pre-processing strategies, and model implementation. This systematic approach ensured consistency in the literature analysis and facilitated robust comparative analysis across all included studies.

III. Results

This section presents the main findings of the 93 studies under review, focusing on the tasks, data characteristics, and methodologies employed. These results provide an overview of the current state of ML applications in MALDI-TOF MS data analysis, highlighting both the strengths and limitations of the approaches used in the field.

A. Key research task

Of the 93 reviewed studies, 19 focus on bacterial species ID [15]–[19], [37]–[39], [41], [47], [57], [61], [63], [76], [77], [90], [94], [96], [114]. While MALDI-TOF MS is a gold-standard technique for bacterial ID, it faces challenges in distinguishing closely related species within the same genus, e.g. *Streptococcus* species, and between species from different genera, e.g. *Escherichia coli* and *Shigella* [31], [115]. These limitations underscore the need for improved methodologies to enhance its discriminatory power. A subset of studies (n=10) address bacterial ID across many genera simultaneously [15], [16], [18], [37]–[39], [41], [60], [61], [77].

AMR is the most common task across the studies, totaling 51 studies focused on predicting resistance profiles to specific antibiotics using ML models applied to MALDI-TOF MS data [20]–[24], [26], [40], [42], [43], [49], [51], [55], [56], [59], [64]–[66], [68], [69], [71]–[75], [78]–[88], [91], [93], [95], [98]–[101], [103], [104], [106]–[112]. Out of these studies,

TABLE I: **Overview of Included Studies (2004–2024).** This table summarizes 93 studies categorized by year of publication. A single quotation mark (’) in the *Isolates* column indicates spectra; otherwise, it refers to bacterial isolates or samples. Algorithms in **bold** represent the primary models used or those achieving the best performance. An asterisk (*) in the *PP* (pre-processing) column indicates steps subjected to ablation studies. Abbreviations for antibiotics, preprocessing steps, peak selection (*PS*) methods, and algorithms are detailed in Tables III and IV.

Year	Paper	Task	Bacteria	Antibiotic	Isolates	PP	PS	Algorithm
2004	Zhang [37]	ID	Bacillus, S.au	-	69	N	-	RBF-**SNN**
	Satten [38]	ID	E.co, S.py, B.an, S.pn	-	120	S	I	**RF**
2008	Hsieh [39]	ID	S.au, E.co, K.pn, S.en, P.ae, GBS	-	132	N	-	QC, SVM, **GA**
2012	Schaumann [40]	AMR	E.co, K.pn, P.ae	ESBL, MBL	76	-	-	**SVM**
2013	Khot [41]	ID	Shigella, E.co	-	138	BaNA	O	**GA**
2014	Nakano [42]	AMR	E.fa	VAN	132	-	GA	**GA**, SNN, QC
	Li [43]	AMR	E.co, K.pn	CTX	103	T	SNR	GA, **SNN**, QC
	Xiao [44]	TYP	M.pn	-	68	BaNA	AUC	**GA**
2015	Zhang [45]	TYP	S.au	-	290	BaSNTA	AUC, SNR	**GA**
	Nakano [46]	TYP	S.pn	-	574	-	-	GA, **SNN**, QC
	Chen [47]	ID	S.pn, S.mi, S.or	-	75	-	S	GA, **SVM**, SNN, QC
2016	Camoez [48]	TYP	S.au	-	95	BaA	-	GA, **SNN**, QC
	Mather [49]	AMR	S.au	VAN	80	MQ	SNR	**SVM**
2017	Ho [50]	TYP	B.fr	-	424	-	AUC	GA, **SVM**, SNN, QC
	Sogawa [51]	AMR	S.au	MET	160	-	-	**SVM**
2018	Wang [52]	TYP	S.au	-	125	BaST	SNR, O	DT, **SVM**, KNN
	Wang [53]	TYP	S.au	-	306	BaS	S, GA	QC, SNN, **KNN**
	Schaumann [40]	ID	Clostridium	-	123	-	-	**SVM**
	Rodrigues [54]	TYP	Klebsiella	-	95	BaSTA	SNR	**SVM**
	Asakura [55]	AMR	S.au	VAN	129	BaSNBi	RF	**RF**
	Wang [56]	AMR	S.au	VAN	125	BaST	RF, SNR	**SVM**, KNN, DT, RF
2019	Ling [57]	ID	Shigella, E.co	-	23	-	-	**SNN**
	Chung [58]	TYP	S.ha	-	254	BaS	RF, AUC	LR, SVM, DT, **RF**
	Tang [59]	AMR	S.au	MET	20	MQ*	RF, O	BI, **RF**
	Wang [60]	TYP	GBS	-	787	BaSTBi*	SNR, O	RF, **SVM**
2020	Desaire [15]	ID	20	-	571	N	O	**AC**
	Pena [28]	TYP	K.pn	-	1,056	BaSTP	SNR, O	**SVM**
	Huang [20]	AMR	K.pn	IPM, MEM	95 ’	Ba	S	SVM, **RF**, KNN, NB, LR
	Ling [61]	ID	14	-	98,000	BaNT	SNR	**SNN**
	Kann [62]	TYP	S.pn	-	69	MQ	O	RF, **CART**
	Wei [63]	ID	B.an, B.ce	-	178	BaNTA	GA, SNR	**GA**
	Weis [26]	AMR	E.co, K.pn, S.au	AMC, CRO, CIP, TZP, PEN	2,676	MQN*	-	LR, PIKE-**GP**
2021	Wang [64]	AMR	S.au	MET	4,858	NBi	S, SNR	DT, **RF**, KNN, SVM
	Gato [65]	AMR	K.pn	MEM, ETP	1,458	BaSA	-	PLS-DA, **RF**
	Mortier [16]	ID	1,000	-	89,767	BaNBi*	-	RF, LR, SVM, **KNN**, **SNN**
	Wang [66]	AMR	E.fa	VAN	7,997	BaSTBi	S	**RF**, SVM, KNN
	Fiamanya [67]	TYP	B.co	-	55	BaSNTP	SNR	**GA**, SNN, **QC**
	Tang [68]	AMR	S.au	MET	960	-	GA	**GA**, SNN, QC
	Chung [69]	AMR	S.au	OXA, CLI, ERY	25,217 ’	BaT*	RF, SNR	DT, SVM, **RF**
	Calderaro [29]	TYP	C.di	-	950	BaSTA	GA, SNR	**GA**, QC, SNN
	Huang [70]	TYP	GBS	-	1,410	BaS	GA	**KNN**, SNN, QC
	Liu [71]	AMR	S.au	MET	452 ’	MQBi	O	RBF-**SVM**, RF
	Calderaro [21]	AMR	E.co	COL	1,040	BaSTAP	S, SNR	GA, QC, **SNN**

TABLE II: **Overview of Included Studies (2004–2024).** This table summarizes 93 studies categorized by year of publication. A single quotation mark (’) in the *Isolates* column indicates spectra; otherwise, it refers to bacterial isolates or samples. Algorithms in **bold** represent the primary models used or those achieving the best performance. An asterisk (*) in the *PP* (pre-processing) column indicates steps subjected to ablation studies. Abbreviations for antibiotics, preprocessing steps, peak selection (*PS*) methods, and algorithms are detailed in Tables III and IV.

Year	Paper	Task	Bacteria	Antibiotic	Isolates	PP	PS	Algorithm
2022	Li [17]	ID	Listeria	-	1,429 ’	BaSN*	O	KNN, NB, RF, SVM, SNN, **DAE-SVM**
	Feucherolles [72]	AMR	C.je, C.co	CIP, ERY, TCY, GEN, KAN, STR, AMP	340	BaNTA	RF, SNR	**RF**, LR, NB
	Weis [22]	AMR	803	42	300,000	MQBi	-	LR, **LGBM**, **MLP**
	Kong [73]	AMR	S.au	MET	548	NBi*	O	GA, **SVM**, DT, RF, PR
	Wang [74]	AMR	E.fa	VAN	7,997	BaST	SNR	**SNN**, RF, XGB
	Wang [75]	AMR	K.pn	CIP	15,782	BaSTBi	S, SNR, O	LR, **SVM**, RF, **XGB**
	Dematheis [76]	ID	B.me, B.ab, B.su	-	44	MQA	RF	**RF**, SVM, SNN, **MNR**
	van den Beld [77]	ID	E.co, S.sp.	-	559	BaSA	O	**SVM**
	Jeon [78]	AMR	S.au	MET	553	BaSNABi	RF	**RF**
	Candela [79]	AMR	E.fa	VAN	178	BaSNA*	-	**SVM**, RF, PLS-DA
	Zhang [80]	AMR	S.au	OXA, CLI	31,815	BaSTBi	SNR, O	LR, **XGB**
	Wang [81]	AMR	K.pn	IPM, MEM, ETP	171	-	RF	RF, SVM, **SVM-k**
	Yu [82]	AMR	S.au	MET	20,359	BaSNBi*	-	**LGBM**, GB, LR, XGB, ET, RF, SVM, DT, KNN, NB
	R.-Villodres [83]	AMR	E.co	TZP	248	BaSNTA*	SNR	PLS-DA, SVM, **RF**
2023	Wang [84]	AMR	S.au	PEN, OXA, ERY, CLI, SXT, FUS	31,807	BaNTBi	SNR, O	**XGB**
	G.-López [85]	AMR	K.pn	15	402	MQ*	O	KSSHIBA-**FA**, RF, SVM, PIKE-GP
	Candela [18]	ID	E.cl	-	337	BaSNA	-	PLS-DA, SVM-L, **SVM-R**, **RF**
	Chung [86]	AMR	E.co	AMX, CAZ, CIP, CRO, CXM	37,918	BaST	RF, SNR	LR, SVM, RF, **XGB**
	Gato [87]	AMR	K.pn	MEM, ETP	715	BaSA*	-	SVM, PLS-DA, KNN, **RF**
	Yu [88]	AMR	K.pn	IPM, MEM, ETP, COL	3,938	BaSNBi	-	**LGBM**
	Abdrabou [30]	TYP	C.di	-	240	BaSNA*	I	SVM, **PLS-DA**, KNN, **RF**
	Zintgraf [89]	TYP	S.pn	-	238	MQAPBi*	O	GLMNET, **RF**, NB, KNN, LDA, SNN, SVM, DT
	R.-Temporal [90]	ID	M.ab	-	325	BaSN	RF	PLS-DA, SVM, **RF**, KNN
	Zeng [23]	AMR	K.pn	IPM	174	BaSNBi	-	**LASSO**, LR, SVM, SNN
	Yu [91]	AMR	K.pn, S.au	IPM, MEM, ETP, MET	170	BaSNBi	-	**LGBM**
	Wang [92]	TYP	S.si	-	205	MQ	-	**RF**, SVM
	Fu [93]	AMR	P.ae	TOB, FEP, MEM	4,039 ’	MQBi	RF	**SNN**
	Manfredi [31]	TYP	E.co	-	202	BaS	S	SNN, **GA-KNN**, QC
	Wang [94]	ID	M.ab	-	102	Bi	RF	**RF**, SVM, DT, LR, KNN
	Iskender [95]	AMR	K.pn	COL	260	BaN	GA	**LDA**, SVM
	Liu [96]	ID	B.lo, B.in	-	100	SN	O	**RF**, SVM, LR
	Candela [97]	TYP	P.ae	-	67	BaSNA	O	PLS-DA, SV, **RF**, NCA-KNN
	Zhang [98]	AMR	K.pn	IPM, MEM, DOR, ETP	2,683	MQBi	O	**SNN**
	Chung [99]	AMR	A.ba, A.no, E.fa, GBS	CIP, VAN, CLI	22,101	MQBaS*	RF	LR, NB, **RF**, SVM
2024	Wang [100]	AMR	S.au	OXA, ERY, CLI, TCY, SXT, MET	5,290 ’	BaSNBi	O	**LGBM**, MLP, RF, KNN, AdaB, LR, SNN
	de Waele [101]	AMR	25	74	55,773 ’	BaSNBi	-	**SNN**, LR, XGB, MLP
	Candela [102]	TYP	C.di	-	69	BaSNA	-	KNN, SVM, PLS-DA, **RF**, **LGBM**
	Lin [103]	AMR	K.pn	CPA	675	NBi	RF	LR, LDA, RF, GB, AdaB, XGB, **LGBM**
	Jian [104]	AMR	K.pn	COL, DOR, IPM	4,307	NBi	O	LDA, **RF**, GB, AdaB, XGB, LGBM, SVM
	Huang [105]	TYP	GBS	-	167	A	GA	GA, SNN, QC, **KNN**
	Godmer [32]	TYP	C.di	-	201	MQ	SNR	MNR, SNN, **RF**, SVM
	Mao [19]	ID	E.co	-	38,400 ’	T	SNR	**LSTM**
	Lin [24]	AMR	S.ma	LVX, SXT	7,266	NBi	RF	LR, LDA, AdaB, **LGBM**, XGB, **RF**, GB
	Xu [106]	AMR	K.pn	ETP, IPM, MEM	240	-	O	**RF**, SVM, LR, XGB
	L.-Cortés [107]	AMR	E.co, K.pn, S.au	CIP, CRO, FEP, TZP, TOB, MEM, FUS, OXA	11,600 ’	BaSBi	-	**SNN**
	Astudillo [108]	AMR	E.co, K.pn, S.au, P.ae	OXA, CLI, FUS, CIP, CRO, TZP, FEP, IPM, MEM	13,157 ’	Bi	-	**MLP**, **SVM**, RF, XGB
	Jian [109]	AMR	K.pn	LVX, CIP	11,996	NBi	O	LR, LDA, **RF**, GB, AdaB, XGB, LGBM
	Macaya [110]	AMR	E.co, K.pn, S.au	CRO, OXA	30,069	NBi	-	**SNN**
	Nguyen [111]	AMR	P.au	11	360	MQBi*	-	LR, RF, **SVM**, LGBM, MLP, **T**
	Ren [112]	AMR	S.ep	20	75,042 ’	MQBi	-	RF, LR, NB, **SVM**, **LGBM**, MLP

TABLE III: Antibiotic codes

Code	Antibiotic	Code	Antibiotic
AMC	Amoxicillin/Clavulanic acid	FUS	Fusidic acid
AMP	Ampicillin	GEN	Gentamicin
AMX	Amoxicillin	IPM	Imipenem
CAZ	Ceftazidime	KAN	Kanamycin
CIP	Ciprofloxacin	LVX	Levofloxacin
CLI	Clindamycin	MBL	Metallo $\beta - lactamases$
COL	Colistin	MEM	Meropenem
CP	Carbapenem	MET	Methicillin
CPA	Ceftazidime-avibactam	OXA	Oxacillin
CRO	Ceftriaxone	PEN	Penicillin
CTX	Cefotaxime	STR	Streptomycin
CXM	Cefuroxime	SXT	Trimethoprim/Sulfamethoxazole
DOR	Doripenem	TCY	Tetracycline
ESBL	Extended spectrum $\beta - lactamases$	TOB	Tobramycin
RY	Erythromycin	TZP	Piperacillin/Tazobactam
ETP	Ertapenem	VAN	Vancomycin
FEP	Cefepime		

TABLE IV: Codes

Code	Pre-processing	Code	Peak Selection
MQ	MALDIquant [113]	AUC	Area Under Curve
Ba	Baseline Subtraction	GA	Genetic Algorithm
S	Smoothing	I	Intensity Thresholding
N	Normalization	RF	Random Forest
T	SNR Thresholding	S	Statistical Tests
A	Averaging	SNR	Signal-to-Noise Ratio Thresholding
Bi	Binning	O	Other
Code	**Algorithm**	**Code**	**Algorithm**
AC	Aristotle Classifier	LGBM	Light GB Machine
AdaB	AdaBoost	LR	Logistic Regression
BI	Binary Discriminant Analysis	LSTM	Long Short-Term Memory
DAE	Denoising Autoencoder	MLP	Multi-Layer Perceptron
DT	Decision Tree	MNR	Multinomial Logistic Regression
CART	Classification Regression Tree	NB	Naïve Bayes
ET	Extra Trees	PLS-DA	Partial Least Squares-Discriminant Analysis
FA	Factor Analysis	PR	Polynomial Regression
GA	Genetic Algorithm	QC	Quick Classifier
GB	Gradient Boosting	RBF	Radial Basis Function
GLMNET	Generalized Linear Model	RF	Random Forest
GP	Gaussian Process	SNN	Supervised Neural Network
KNN	k-Nearest Neighbours	SVM	Support Vector Machine
LASSO	Least Absolute Shrinkage and Selection Operator	T	Transformer
LDA	Linear Discriminant Analysis	XGB	Extreme GB

some (n=10) also include bacterial species identification, as they predict AMR for more than one bacterial species [22], [26], [40], [43], [91], [99], [101], [107], [108], [110].

Of the examined studies, n=23 focus on TYP, involving only GT [28], [67], [97], only RT [29], [30], [32], only ST [31], [89], [92], or multiple TYP methods [62], [70]. As an example of a TYP study, Huang et al. [70] focus on predicting the main ST of Group B *Streptococcus*, where some of them are found to be strongly associated with neonatal invasive diseases.

While most studies focus exclusively on a single task, only one paper adopts a multi-task approach, simultaneously predicting both ID and AMR [101]. Among papers performing AMR detection for multiple bacterial species, AMR prediction is always carried out separately for each species. Similarly, studies predicting resistance to multiple antibiotics generally train one model per bacteria-antibiotic pair. Only n=2 papers perform species-specific AMR prediction for all tested antibiotics [101], [108]. This limited adoption of multi-task approaches underscores the untapped potential of ML algorithms in addressing multiple related tasks simultaneously. Up to date, no existing study addresses all three tasks or combines both ID and TYP simultaneously.

B. Data acquisition device

The reviewed studies employ various MALDI-TOF MS data acquisition devices. The two main manufacturers of these instruments are Bruker Daltonics [116] and BioMérieux [117]. Both provide a full pipeline, from performing MALDI-TOF MS measurements of the cultured isolate to species identification based on database lookup. Nevertheless, species identification from MALDI-TOF mass spectra remains an active field of research, as several bacteria, such as *Burkholderia cepacia* complex or *Citrobacter freundii* complex, are not identifiable at the species level [26].

The most common instrument is the Bruker Microflex Biotyper (n=74), which is favored due to its high throughput capabilities and well-established software tools, such as flexAnalysis [118] and ClinProTools [119], which facilitate data analysis and integration with ML algorithms. In comparison, the VITEK MS system by BioMérieux (n=10) appears less frequently in the reviewed literature. VITEK MS is valued for its ease of use in clinical settings but lacks the widespread adoption of research seen with Bruker systems. One study [100] uses data acquired from Bruker as well as bioMérieux.

In addition to these two primary devices, a few studies (n=8) make use of less common MALDI-TOF instruments or customized systems for specific research purposes, such as, Zybio Inc., Applied Biosystems, PerSeptive Biosystems, Absciex, ASTA MicroIDSys, or Voyager DE-STR.

C. Data size

Distinguishing between bacterial samples or isolates and mass spectra is essential in the context of MALDI-TOF MS studies. A bacterial sample or isolate refers to a cultured strain obtained from a clinical or environmental source, representing a unique biological entity. In contrast, a spectrum represents the protein fingerprint of an isolate, generated through MALDI-TOF MS analysis. Depending on the experimental design, a single isolate may yield multiple spectra due to biological or technical replicates or varying acquisition or preprocessing conditions.

This distinction is critical for interpreting dataset sizes, as some studies describe their datasets in terms of the number of samples or isolates [20], while others provide the number of spectra generated in the MALDI-TOF MS analysis [61], leading to potential discrepancies. Others provide the number of isolates and the number of technical replicates they acquire [30], being able to quantify how many spectra are available. In general, isolates are the most relevant unit of measurement for ML applications, but some studies (n=13) only report the number of MALDI-TOF spectra, leaving the number of isolates unspecified. This lack of standardization in reporting data sizes complicates comparisons across studies. For clarity, Table II lists the number of isolates under the column *Isolates*, with an ' indicates that the number represents spectra when isolate data cannot be inferred from the original study.

The size of the datasets used in the reviewed studies varies significantly. For example, Pena et al. [28] analyzed just 44 bacterial isolates, whereas Weis et al. [22] utilized a massive dataset of over 300,000 isolates. Among the studies, the

majority (n=57) rely on relatively small datasets with fewer than 500 isolates, while a low number of papers (n=14) employed medium-sized datasets ranging from 500 to 10,000 isolates. Only a few studies (n=9) incorporated very large datasets exceeding 10,000 samples. The median dataset size among the reviewed studies is 209 isolates and 1,443.5 spectra. Smaller datasets, while easier to manage, often compromise the reliability and generalizability of models, particularly in the absence of robust external evaluation. The scarcity of large-scale datasets poses a significant challenge, underscoring the need for further research and development to address this limitation.

Data balance is critical for minimizing biases in ML models, yet fewer than one-third of the studies reviewed (n=30) report working with balanced datasets [20], [21], [23], [64], [66], [68], [69], [71], [74], [78], [79], [81]–[83], [87], [94], [95], [99]. Conversely, data imbalance is prevalent (n=53), particularly where certain antibiotic resistance profiles in AMR tasks or strains in TYP are underrepresented. Alarmingly, the remaining studies (n=10) did not provide information on whether their datasets were balanced, highlighting a significant gap in reporting practices.

To address this challenge, several studies (n=16) applied *resampling* techniques, using oversampling and/or undersampling to balance minority and majority classes [24], [26], [80], [85], [103], [106], [109]. A specific form of oversampling, the Synthetic Minority Oversampling Technique (SMOTE), was utilized in some studies (n=5) [75], [93], [110], [111], while Tomek Links, an undersampling approach, was applied in one study [75]. Some studies (n=15) indicate addressing this issue by making use of data imbalance-aware evaluation metrics such as the area under the precision-recall curve (AUPRC) or balanced accuracy, as in [22], [26], [72], [85], [98], [100], [107], [111], [112]. Additionally, few studies (n=3) tackled imbalance by weighting the loss function during model training [26], [84], [110], [111].

However, other studies (n=27) acknowledged working with unbalanced datasets without proposing any specific balancing techniques [62], [65], [77], [86], [88], [91], [97].

In summary, while data description is a critical factor in the development of robust ML models, the lack of standardization in how data is reported—whether by samples, isolates, or spectra—along with issues like data imbalance, can limit the comparability and effectiveness of the models across different studies.

D. Bacteria

The reviewed studies primarily focus on a few key bacterial species that are clinically significant due to their association with hospital-acquired infections and AMR. The most prevalent species studied is *Staphylococcus aureus* (n=26) is extensively studied, particularly in the context of *methicillin-resistant S. aureus* (MRSA), a notorious pathogen responsible for numerous healthcare-associated infections [120], [121].

Following closely, *Klebsiella pneumoniae* (n=21), reflecting its critical role in carbapenem-resistant infections, which pose a major challenge in healthcare settings [122]. *Escherichia coli* (n=16), known for its role in urinary tract infections and rising resistance to antibiotics such as fluoroquinolones [123] and carbapenems [124], is also a frequent subject of analysis.

The majority of studies (n=73) focus on bacteria from a single genus, such as *Streptococcus* [62], [70], [89], [92]. In contrast, a smaller subset (n=20) examines bacteria across multiple genera, of which only a few (n=7) addressing more than three genera [15], [16], [22], [61]. This tendency likely stems from the focus of most AMR-related studies on single-species bacterial analysis for susceptibility to one or multiple antibiotics (see Section III-E).

In summary, *S. aureus* (n=26), *K. pneumoniae* (n=21), and *E. coli* (n=16) are the most studied bacterial species. While multi-bacteria studies are becoming more prevalent, addressing the growing need for diagnostic tools capable of handling a wider range of pathogens, they still represent only about 21% (n=20/93) of the reviewed studies.

E. Antibiotics under study

The reviewed studies investigate a broad range of antibiotics to predict AMR across bacterial species; a total of 28 different antibiotics are studied. The number of antibiotics analyzed per study varies widely, with some focusing on a single antibiotic, while others assess resistance to a broader spectrum. The maximum number of antibiotics studied in a single paper reaches 42 [22], highlighting the versatility of ML models in addressing complex, multi-drug resistance profiles. However, there is a clear pattern where studies either focus on a single bacteria-antibiotic pair (n=15) [21], [23], [64], [68], [71], [73]–[75], [78], [79], [81]–[83], [95], [103] or attempt a more comprehensive analysis involving multiple antibiotics (n=18) [20], [65], [69], [80], [81], [84]–[88], [93], [98], [100], [103], [106], [109], [111], [112]. The less frequent case is the multiple bacteria and multiple antibiotics (n=7) study [22], [26], [72], [91], [99], [107], [110].

Among the antibiotics studied, the most frequently analyzed (n=13) are *carbapenems* (imipenem, IPM; meropenem, MEM; ertapenem, ETP; doripenem, DOR) [20], [23], [65], [81], [85], [87], [88], [91], [93], [98], [106], [107], [109]. This is due to their critical role as antimicrobial agents for infections caused by multidrug-resistant pathogens. However, the emergence of carbapenem-resistant bacteria has increased dramatically posing a serious public health threat [125], which makes them a primary focus for AMR research.

In addition to carbapenems, other commonly studied antibiotic groups include:

- *Beta-lactams*, such as methicillin (MET) (n=8) and penicillins (PEN) (n=2), especially in the context of MRSA form [64], [68], [71], [73], [78], [82], [88].
- *Glycopeptides*, with vancomycin (VAN) (n=4) being central in studies of *Enterococcus faecium*, particularly for vancomycin-resistant *Enterococci* (VRE) [66], [74], [79], [99].
- *Fluoroquinolones*, such as ciprofloxacin (CIP) (n=6), commonly analyzed in *E. coli* and *K. pneumoniae* for urinary tract and systemic infections [26], [75], [107].

- *Polymyxins*, notably colistin (COL) (n=4), which is another last-resort antibiotic frequently studied for resistance in *K. pneumoniae* and other Gram-negative bacteria [21], [88], [95], [109].

These studies also examine other groups of antibiotics, such as *cephalosporins*, e.g., ceftriaxone (CRO) (n=4) [107], ceftazidime (CAZ) (n=1) [86]; and *aminoglycosides*, e.g., gentamicin (GEN) (n=1) [72], tobramycin (TOB) (n=2) [93], to capture a wider range of resistance mechanisms. The variety of antibiotics studied reflects the complexity of bacterial resistance and the need for comprehensive profiling to inform treatment strategies.

In summary, the most common antibiotics studied include carbapenems (n=13), MET (n=8), and CIP (n=6). Carbapenem resistance, in particular, is a frequent research focus due to the critical role of these drugs in treating multidrug-resistant infections. Several studies also investigate other antibiotic groups, aiming for a deeper understanding of multi-drug resistance patterns and further demonstrating the potential of ML in AMR prediction. For further details and a comprehensive overview of the antibiotics used in each study, please refer to Table II.

F. Pre-processing

Pre-processing is a crucial step in the analysis of MALDI-TOF MS data, as it directly influences the performance of ML models. The concept of "garbage in, garbage out" [126] applies strongly here, emphasizing the need for standardized processes to ensure that high-quality data is fed into algorithms. Spectrometry data typically contains several forms of noise, such as baseline shifts or fluctuations, which must be corrected to retain only relevant information. Moreover, not every part of the mass spectrum is informative; some regions contain valuable signals, while others do not contribute to the classification task, making techniques like peak selection and signal-to-noise ratio (SNR) thresholding essential for focusing on the most relevant parts of the spectrum. Normalization is also crucial, because of the wide range of intensity values in mass spectra, which can vary by thousands of units. Additionally, spectra may be acquired from different devices (see Section III-B), each with unique characteristics, requiring standardization across machines to ensure that the models are robust and generalizable. Furthermore, MALDI-TOF MS data can vary in length, while most ML models require a fixed input size. Techniques like binning or interpolation are often applied to create uniform input sizes that models can process. Finally, technical replicates—multiple spectra generated from the same sample—are commonly collected and should be averaged to avoid duplicated data that could introduce bias. These steps are essential for ensuring data integrity and maximizing model performance.

While open-source (MALDIquant [113]) and commercial software (ClinProTools [119]) are available to process and analyze MALDI-TOF MS, standard pre-processing involves several steps with numerous parameter choices, which are often not justified. The inconsistency in the choice and order of pre-processing steps presents a significant challenge in the field, as it complicates the comparison of results across studies and hinders reproducibility. While MALDIquant [113] is a ready-to-use R package designed for MS data pre-processing, only some of the reviewed papers adopt this tool (n=14; see Table II). Even among these, the package's extensive customization options lead to considerable variability, as researchers can adjust hyperparameters and tailor steps like smoothing, baseline correction, normalization, and peak detection.

Although MALDIquant provides a standardization framework, the flexibility it offers often results in unique pre-processing pipelines across studies. For instance, Weis et al. [26] employed a pipeline comprising (i) normalization, (ii) Savitzky-Golay smoothing, (iii) baseline correction, (iv) Total Ion Current (TIC) normalization, (v) SNR thresholding, (vi) warping, and (vii) trimming. In contrast, Dematheis et al. [76] used a different sequence: (i) spectral cleaning, (ii) transformation, (iii) Savitzky-Golay smoothing, (iii) baseline correction, and (iv) normalization. These variations illustrate how methodological choices in pre-processing can impact data analysis and interpretation.

Similarly, studies that use ClinProTools [116] (n=7) [21], [29], [31], [63], [67], [68], [70] do not follow a standardized pre-processing pipeline. However, several common pre-processing steps are consistently applied across different studies. These include baseline subtraction (n=37), smoothing (n=30; e.g., Savitzky-Golay), normalization (n=26; e.g., TIC), SNR thresholding (n=17), spectra averaging (n=15), binning (n=22), and principal component analysis (PCA) for dimensionality reduction (n=13), see column *Pre-processing* of Table II. These eight steps are the most common among the studies reviewed, but there are many others that, for simplification, are not specified here. Additional pre-processing steps include scaling, spectra and peak alignment, recalibration, binarization, and various transformations. For more details on the specific order of steps applied, please refer to the individual papers. Table II provides a simplified overview of the principal steps applied and whether a study performs any ablation or comparative studies on these steps, indicated by an asterisk (*). Although ablation studies offer valuable insights, they are not yet widespread enough to establish clear recommendations or a standardized approach.

G. Algorithms

A variety of ML algorithms is applied across the studies reviewed. While most studies (n=49) compare algorithm performance (e.g., [16], [17], [20], [24], [82], [89], [109]), the choice of algorithm is generally guided by dataset size, quality, and computational resources rather than the specific task. Notably, minimal correlation exists between the selected model and primary tasks such as ID, AMR, and TYP.

In clinical settings, the preference often leans towards simple and interpretable algorithms to ensure explainability. Classical ML models like Support Vector Machines (SVM) are widely used for their robustness and capacity to handle high-dimensional data. While Random Forests (RF) are particularly popular (n=20, see Table II) due to its bagging structure that improves predictive accuracy and mitigates overfitting.

Moreover, RF is frequently employed as a feature selection tool (n=7) within study workflows to reduce dimensionality and focus on relevant peaks [24], [69], [72], [76], [81], [93], [112]. Other feature engineering techniques include logistic regression (LR) [109], PCA [63], LASSO [71], and decision trees (DT) [73].

SVMs (n=13) are another preferred choice, particularly for classification tasks involving high-dimensional datasets where the number of features far exceeds the sample size [69], [71], [73], [75], [79], [81], [111], [112]. Boosting methods such as gradient boosting (GB), adaptive boosting (AdaB) [127], light gradient boosting (LGBM), and extreme gradient boosting (XGB) [128] also show strong performance (n=10). Discriminant analysis methods, including partial least squares-discriminant analysis (PLS-DA) and linear discriminant analysis (LDA) (n=4), are occasionally applied. Probabilistic Bayesian approaches, such as Gaussian processes (GP) and Factor Analysis (FA) approaches [129], are less commonly used (n=2). Additional classical ML methods include k-nearest neighbors (KNN) (n=4) and genetic algorithms (GA) (n=6).

Classical ML methods dominate over Depp Learning (DL) approaches (n=10), including MultiLayer Perceptrons (MLP), Dense Neural Networks (DNN), Convolutional Neural Networks (CNN), AutoEncoders (AE), and Long Short-Term Memory Neural Networks (LSTM-NN). NNs are particularly useful with raw spectra, as they excel at extracting local patterns and can accommodate the large lengths typical of mass spectra. NNs are shown to outperform other ML models when trained on large datasets, i.e., several thousand spectra [16], [61], [74], [107]. Conversely, when data availability is limited, classical ML methods often achieve the best results [29], [68], [76]. DL approaches have been explored in only one study [111], where a transformer model was trained for ID and subsequently used in a transfer learning pipeline for AMR prediction, leveraging an SVM for the downstream task.

Some studies also experiment with ensemble methods (n=3) [95], [99], [111] and hybrid models (n=4) that combine multiple algorithms to improve accuracy and generalizability [17], [31], [32], [70]. For example, Li et al. [17] identified a Denoising AE-SVM ensemble as the best-performing model in their multi-combination ablation study. This approach leverages the strengths of individual algorithms to yield more robust results.

In summary, while there is no one-size-fits-all solution, RF, SVM, and boosting models are the most well-performing algorithms (n=50) for MALDI-TOF MS data analysis. Ensemble approaches and algorithm comparisons are increasingly employed to enhance predictive accuracy. DL and NNs approaches are gaining relevance (n=10).

In summary, while there is no one-size-fits-all solution, RF, SVM, and boosting models remain the most effective and widely used approaches (n=50) for MALDI-TOF MS data analysis. Ensemble methods and algorithm benchmarking are increasingly applied to improve predictive accuracy. DL and NNs, although less frequently used (n=10), are gradually gaining traction in this domain.

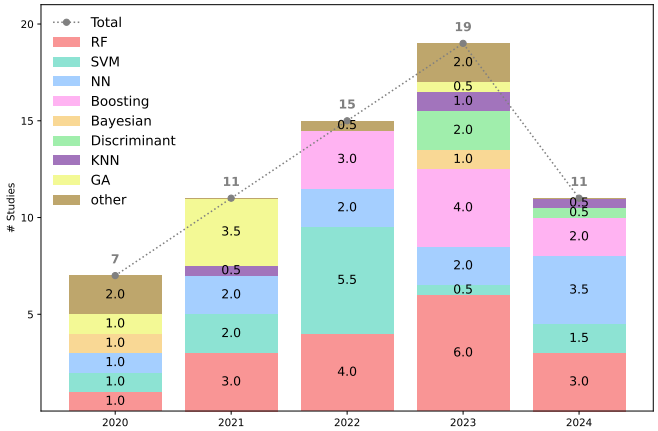

Fig. 3: Papers per year and their respective principal model.

H. Software

The majority of studies reviewed utilize Python (n=32) [130] as the primary software for analyzing MALDI-TOF MS data. Python's popularity is driven by its versatility, extensive ML libraries (e.g., scikit-learn [131], TensorFlow [132], XGBoost [128]), and ease of integration with pre-processing and analysis tools. Its flexibility and widespread adoption enable the development of reproducible, open-source workflows, which are critical for advancing standardization in the field.

R [133] is used in a smaller subset of studies (n=10) [64], [66], [75], [76], [94], particularly those employing MALDIquant for pre-processing (n=5) [32], [62], [71], [76], [89]. R's strengths in statistical analysis and specialized mass spectrometry tools make it a valuable resource, though its usage is less common than Python. MATLAB [134], a proprietary software, is used in only one study [95], valued for its robust matrix computations and visualization capabilities, though its high cost and closed nature limit broader adoption.

Manufacturer-specific software is employed in several studies (n=9), with Bruker's ClinProTools being the most prevalent (n=7) [21], [29], [31], [63], [67], [68], [70], [116], [119]. Other proprietary tools include Bionumerics (n=1) [77], [135] and AMRQuest (n=1) [78], [136], tailored specifically for MALDI-TOF MS data analysis.

Clover [137], a web-based MALDI-TOF MS analysis platform, is featured in eight studies [18], [30], [65], [79], [83], [87], [90], [97], and is particularly popular in clinical settings. Designed for accessibility, Clover requires minimal data science expertise, offering a streamlined workflow with built-in pre-processing and ML models. However, its proprietary nature limits flexibility compared to Python-based solutions, which allow for greater customization and adaptability.

Given the increasing complexity of MALDI-TOF MS data and the need for adaptable workflows, Python remains the standard for future studies. Its extensive library ecosystem, strong community support, and open-source ethos make it ideal for developing highly reproducible and flexible analytical workflows.

I. Cross-validation

Cross-validation (CV) is widely used across most of the studies reviewed (n=56), serving as a critical step to ensure the robustness and generalizability of ML models. This technique evaluates model performance by partitioning the dataset into k equal parts, training on k-1 parts while using the remaining fold for testing. This process is repeated k times, with each fold serving as the test set once, providing a comprehensive assessment of the model's performance. Five-fold and ten-fold CV are the most frequently used variations, balancing computational efficiency and model reliability, and k-fold CV is adopted by most studies (n=55).

A less common variant is leave-one-out cross-validation (LOO-CV) (n=1), used primarily in cases where datasets are small, as in Huang et al. [20], where only 95 isolates were available. In LOO-CV, each data point is used as a test set exactly once, while the model is trained on all other points. Though computationally demanding, this approach is exhaustive and ensures each data point is individually validated, making it particularly suited for smaller datasets.

When using CV, particularly with large datasets that may include technical replicates, it is crucial to stratify k-fold splits by both, class and species, to prevent information leakage between training and testing sets. Weis et al. [22] emphasize this need, noting that failing to do so can introduce bias and compromise the validity of results. Although most studies default to k-fold CV, few explicitly address the risk of improper data splitting which may be leading to data leakage and post-split quality checks are rarely conducted. While CV can enhance model performance, its misuse may lead to misleading outcomes. Mortier et al. [16] highlight that CV is a questionable approach, especially when many technical replicates are present.

J. Evaluation metrics

The most commonly used evaluation metrics in the reviewed studies are accuracy, sensitivity (recall), specificity, and the area under the curve (AUC) of the receiver operating characteristic (ROC) curve. Accuracy is widely included (n=37) as a measure of overall model performance. It indicates the proportion of correctly classified instances out of the total. However, due to potential class imbalances in many datasets, accuracy alone can be misleading, particularly when the number of resistant samples is disproportionately low in AMR or the number of samples per genus and species are imbalanced in ID. To address this, sensitivity (n=30) and specificity (n=30) are also frequently reported, measuring the model's ability to correctly identify positive and negative cases, respectively.

Precision (n=4) and F1-score (n=10) are also used, with the F1-score being particularly valuable for imbalanced datasets as it combines precision and recall into a single metric. Precision indicates the proportion of true positive predictions among all positive predictions, while the F1-score balances both precision and recall, making it a useful metric in situations where both false positives and false negatives carry significant weight.

The area under the receiver operating characteristic AUC-ROC curve (n=39) and AUPRC (n=12) are increasingly used metrics. These metrics are commonly applied regardless of the specific task in hand, indicating a trend toward evaluating model performance with more sophisticated metrics, even when their relevance may vary by application. AUC-ROC provides a measure of a model's overall discriminative ability across all thresholds, making it useful in binary classification tasks, though it may be less informative for imbalanced datasets. AUPRC, on the other hand, focuses specifically on the positive class, providing a more detailed performance evaluation when the minority class (e.g., resistant bacteria) is of particular interest. AUPRC, as mentioned in section III-C, is commonly used as a data imbalance-aware evaluation metric. The use of both metrics suggests an awareness of the limitations of simpler metrics like accuracy, yet there is little indication that these choices are tailored to the specific demands of each of the reviewed tasks.

Less frequently, metrics like Jaccard (n=1), negative predictive value (NPV) (n=8), positive predictive value (PPV) (n=9), Matthews correlation coefficient (MCC) (n=2), and Kappa (n=2), used to further evaluate not only overall performance but also the model´s ability to correctly predict both minority and majority classes.

K. External validation

External validation refers to the process of evaluating a ML model using an *independent* dataset that was not involved in the model's training or internal validation phases. This technique is crucial for assessing the model's ability to generalize beyond the specific data on which it was developed, proving its robustness when applied to new or unseen distributions of data. However, despite its importance, external validation is often misunderstood or confused with standard internal testing procedures like CV, where the same dataset is partitioned into training and testing subsets. True external validation involves testing the model on completely independent datasets, which many studies fail to implement.

In the reviewed studies, several (n=24) conduct proper external validation [16], [18], [19], [22], [61], [64]–[67], [69], [74], [80], [82], [84], [86], [87], [89], [90], [96], [98], [99], [107], [111], [112]. The remaining studies rely on internal validation methods such as CV or standard train-test splits, which, while effective for model tuning, do not fully demonstrate the model's generalizability to other environments. The lack of external validation raises concerns about how well these models will perform when applied in real-world scenarios, where data can vary significantly due to differences in acquisition devices, settings, or sample populations. External validation is critical for proving that a model is not overfitting to the specific dataset it was trained on and is an essential step for models intended for clinical use or broader application.

L. Open-source data and code

The availability of open-source data and code is essential for advancing research, promoting transparency, and ensuring reproducibility in ML studies. However, relatively few studies

TABLE V: Open-source studies

Paper	Database	Size	Form	Code	Software	Comments
Desaire (2020) [15]	✓	571	raw			UCI
Weis (2020) [26]				✓	Python	GitHub
Weis (2022) [22]	✓	132,097	raw, pp	✓	Python	DRYAD used in [93], [107], Github
Wang (2022) [74]				✓	Python	GitHub
Yu (2022) [82]	✓	2,456	pp			Zenodo
Wang (2023) [84]	✓	31,807	pp	✓	R	Github
G.-López (2023) [85]	✓	2,000	raw, pp	✓	Python	Github
Candela (2023) [18]	✓	1,566	raw			Clover
Chung (2023) [86]	✓	37,918	pp			Github
Yu (2023) [88]	✓	1,853	pp			Zenodo
Zintgraff (2023) [89]	✓	193	raw	✓	R	Mendeley
Wang (2023) [92]	✓	205	pp	✓	Python	Supplemental material
Mao (2024) [19]	✓	48	?			SimTK
L.-Cortés (2024) [107]				✓	Python	Github
Nguyen (2024) [111]	✓	30	raw	✓	Python	Github
Ren (2024) [112]				✓	Python	Github
RKI dataset [34]	✓	11,038	raw			Zenodo
Sum:	**12**	218,115		**10**		

in our review provide public access to their datasets or make their code available on platforms like GitHub [138] or other public repositories. This lack of *open-sourceness* restricts other researchers' ability to replicate findings, build upon existing work, and apply models to new datasets.

Among the reviewed studies, only a small proportion share either the dataset [15], [18], [19], [82], [86], [88] or the code [26], [74], [107]. Even fewer studies provide access to both [22], [84], [89], [92], limiting the broader scientific community's ability to fully evaluate the methods and results. Table V summarizes which studies have publicly available code or data. While some studies make their models or pipelines open-access, the underlying data—especially sensitive clinical data—often remains restricted due to privacy concerns or institutional policies. This lack of access makes validating models on external datasets difficult, as independent researchers cannot readily replicate the analyses.

Two large public MALDI-TOF MS datasets are available to the research community: the RKI [34] and the DRIAMS [33] dataset published by Weis et al. [22]. Both provide a comprehensive collection of spectra from various bacterial species, enabling large-scale studies on microbial identification and antimicrobial resistance (AMR). These repositories have not only facilitated the replication and validation of existing models but have also spurred new research. For instance, López et al. (2024) [107] and Fu et al. (2023) [93] utilized the DRIAMS dataset to develop novel ML models, illustrating the transformative impact of open-access data on advancing the field.

Despite the growing availability of MALDI-TOF MS datasets, only two contain data labeled for serotyping [89], [92]. The absence of open-access serotyping data restricts researchers' ability to develop and validate models for hypervirulence prediction across bacterial strains. Conversely, six of the eleven public datasets contain AMR-labeled samples for multiple antibiotics. Only Yu et al. [91] focuses on a single antibiotic, Methicillin, while the other AMR datasets [22], [34], [84], [86], [88] involve at least four different antibiotics. The most comprehensive dataset, DRIAMS [22], [33], includes over 100,000 spectra and covers 42 different antibiotics.

The importance of open-source science cannot be overstated. When databases and code are made publicly accessible, it fosters a collaborative environment where researchers worldwide can contribute, refine, and adapt models for diverse applications. Open science democratizes access to advanced tools and methodologies, ensuring that progress is not confined to a few institutions but benefits the broader research and medical communities. Embracing open-source practices is vital for the future of ML in MALDI-TOF MS data analysis, as it promotes transparency, accelerates discovery, and supports science that is accessible to all.

However, standardization is essential for reproducibility. As shown in Table V, very few datasets are shared in their original file format (e.g., BrukerRaw) [18], [34], [89], which retains valuable metadata such as acquisition time and device information. Additionally, data is often published after preprocessing, as seen in [84], [86], [92], limiting its usability for other research purposes. Wherever possible, data should be shared in raw form with accompanying metadata, enabling diverse applications and maximizing the potential of open science.

IV. Discussion

The application of ML to MALDI-TOF MS has demonstrated significant potential in clinical microbiology, enabling faster and more accurate bacterial ID, AMR prediction, and TYP. However, this review highlights that the majority of studies remain task-specific, often addressing these challenges in isolation. While classical ML algorithms like RF and SVM dominate the landscape due to their simplicity and robustness, the emerging use of DL models points to a promising avenue for handling complex, high-dimensional spectral data. Nevertheless, the inconsistent application of pre-processing methods, limited exploration of multi-task approaches, and reliance on black-box tools present obstacles to both reproducibility and clinical trust. Furthermore, the lack of standardized evaluation metrics and external validation undermines the generalizability of findings, particularly in real-world scenarios where diverse bacterial populations and clinical environments must be considered.

To ensure broader adoption and facilitate clinical translation, addressing these gaps is essential. This involves implementing reproducible pre-processing workflows, exploring multi-task and transfer learning frameworks, and promoting open-source datasets and code to enhance transparency and collaboration. Additionally, the integration of explainable models and robust evaluation protocols will help bridge the gap between microbiology and ML, fostering trust among clinicians and researchers. The following guidelines build on these findings, offering actionable recommendations to standardize methodologies, encourage interdisciplinary collaboration, and drive innovation in the application of ML to MALDI-TOF MS.

A. Guidelines

After reviewing X state-of-the-art papers, we identified a lack of standardization in ML research on MALDI-TOF MS data. Based on our expertise in ML, we propose the following guidelines to ensure a well-structured research pipeline that can effectively advance ML applications in this field.

It is essential to clearly define the research question, specifying how MALDI-TOF MS data will be used and how ML techniques will contribute to answering the question. Once the question is established, the pipeline should proceed through the following stages: (i) Exploratory Data Analysis (EDA), (ii) pre-processing, (iii) model selection, and (iv) model evaluation.

In the EDA phase, we must focus on understanding the dataset to decide which preprocess and model to use. Key considerations include (i) documenting the number of unique bacterial isolates, which will guide model selection; (ii) number of replicates, which will condition pre-processing techniques; and (iii) data imbalance, which will condition pre-processing, losses used and evaluation methods.

Pre-processing standardizes the dataset to improve model learning. While the SOTA lacks consistency, we propose the following mandatory steps for all raw spectra: (i) variance stabilization by log2 or square root transformation, (ii) baseline removal using methods like SNIP, top-hat, or ALS, (iii) smoothing with a Savitzky-Golay filter, and (iv) trimming between 2000 and 20000 Da, to remove noise. After these steps, we suggest either working with different-length spectra or standardizing their length. As baseline method, for simplicity and extension in the SOTA, we suggest binning with a 3 Da bin size, converting spectra into a 6000-element vector by summing intensities in each bin, making the data easier to use in ML/DL models. For ML experts, varying lengths can be used with models such as Transformers [111] or PIKE [26]. For handling replicates, we suggest two approaches: (1) averaging spectra to create a unique representation per isolate, and (2) treating replicates as a multi-instance problem, ensuring replicates from the same isolate are in the same split to avoid data leakage. If using your own implementations, we recommend the second approach, as it leverages more data and is robust to reproducibility issues. For black-box software, averaging replicates is simpler. To address data imbalance, we recommend two strategies: (1) oversampling with techniques such as SMOTE [110], [111], and (2) using imbalanced-aware loss functions such as weighted Log-Loss in LR, weighted Hinge Loss in SVM, or class weights in Random Forest models.

Regarding model selection, the goal is to select the best model that fits both the research question and the available data. For small datasets ($N \ll D$), classical ML models like RF (n=20) or LGBMs (n=10) are generally recommended, as they serve well for hypothesis testing. Hyperparameter tuning should always be done using inner cross-validation (e.g., 3-fold or 5-fold) and clearly documented. Likewise, Bayesian models, like GPs [26], are ideal for small datasets because they generalize better and do not require an inner CV. Alternatives like DBLFS for LR or Factor Analysis [85] for combining multimodal data are also useful. Larger datasets, e.g, $N \gg D$, open up the possibility of using complex and accurate DL models like MLPs [22] or Transformers [111] are better suited to handle large datasets and complex representations of MALDI-TOF spectra. Unsupervised learning methods such as PCA, t-SNE, and UMAP are useful for visualization and interpretability, but should not be used to solve classification tasks (ID, TYP, AMR), which are better addressed by supervised models.

Evaluation must include both internal and external validation. Use outer cross-validation (CV) by splitting the dataset into k folds (typically 3, 5, or 10), training and testing the model k times. Calculate the mean and standard deviation of evaluation metrics to determine the most robust model. Within each fold, apply inner CV to tune hyperparameters. After selecting the best model, retrain it using the entire internal dataset. External validation, also denoted as OOD, is crucial and should involve a completely separate dataset (e.g., from a different MALDI-TOF machine or hospital, to ensure OOD) to confirm the model's generalizability to unseen data distributions. Evaluation metrics should consider replicate handling by ensuring all replicates of the same isolate stay within the same split. For imbalanced datasets, use balanced accuracy instead of standard accuracy, and consider metrics like AUC-ROC, precision, recall, and AUPRC to account for class imbalance in AMR, ID, and TYP tasks.

V. Conclusions

This review systematically examined the integration of ML with MALDI-TOF MS for bacterial ID, AMR prediction, and strain TYP, covering 93 studies published between 2004 and 2024. The findings reveal the transformative potential of ML in clinical microbiology, particularly in enhancing diagnostic speed, precision, and scalability. Classical ML algorithms, such as RF and SVM, remain the most widely adopted, offering robust performance for small to medium-sized datasets. Recent advancements in DL models, though less commonly applied, show promise in handling high-dimensional mass spectral data and addressing complex tasks.

Despite significant progress, this review identified several limitations in current research. A lack of standardized preprocessing pipelines, inconsistent reporting practices, and reliance on black-box models hinder transparency, reproducibility, and clinical adoption. Moreover, the limited availability of large, high-quality, open-access datasets restricts the development of generalizable models. While most studies address singular tasks like ID or AMR prediction, the potential of multitask approaches remains largely untapped. Only two papers demonstrated species-specific AMR prediction for multiple antibiotics, highlighting an area for future exploration.

VI. Future Directions

To address these gaps and propel the field forward, future research should focus on the following directions:

- **Standardization of pre-processing pipelines:** The development of standardized, reproducible pre-processing workflows is critical for ensuring consistency across studies. Open-source tools such as MALDIquant should be leveraged, and ablation studies should become standard practice to assess the impact of each pre-processing step.
- **Multi-task and Transfer Learning:** The application of multi-task learning approaches to simultaneously tackle ID, AMR, and TYP could enhance model versatility and efficiency. Transfer learning, particularly with pre-trained DL models, should also be explored to maximize the utility of existing datasets.
- **Open-Access Data and Code:** Expanding public repositories, such as DRIAMS and RKI, with annotated data for TYP and AMR tasks is essential. Open-sourcing datasets and code will foster collaboration, enhance reproducibility, and accelerate innovation in the field.
- **Addressing Data Imbalance:** Future studies should employ advanced data augmentation techniques, such as generative models and synthetic data creation, to address class imbalances in both ID and AMR datasets. Metrics like AUPRC and balanced accuracy should replace accuracy in imbalanced scenarios.
- **Explainability and Validation:** The development of interpretable ML models and rigorous validation practices, including external validation and LOO-CV, are essential for clinical adoption. Emphasizing model transparency will build trust and facilitate integration into real-world workflows.
- **Integration with Emerging Technologies:** Advancements in generative AI, federated learning, and real-time mass spectral analysis offer new opportunities for improving performance, scalability, and privacy. Future research should explore the potential of these technologies to further advance MALDI-TOF MS applications.

By addressing these challenges and embracing interdisciplinary collaboration, the integration of ML with MALDI-TOF MS has the potential to revolutionize clinical microbiology, offering faster, more precise, and scalable diagnostic solutions for a wide range of pathogens and resistance profiles.

References

[1] F. Hillenkamp, M. Karas, R. C. Beavis, and B. T. Chait, "Matrix-assisted laser desorption/ionization mass spectrometry of biopolymers," *Analytical Chemistry*, vol. 57, no. 14, pp. 2299–2304, 1985.

[2] A. Tran, K. Alby, A. Kerr, M. Jones, and P. H. Gilligan, "Cost savings realized by implementation of routine microbiological identification by matrix-assisted laser desorption ionization–time of flight mass spectrometry," *Journal of clinical microbiology*, vol. 53, no. 8, pp. 2473–2479, 2015.

[3] F. Nomura, "Proteome-based bacterial identification using matrix-assisted laser desorption ionization–time of flight mass spectrometry (maldi-tof ms): A revolutionary shift in clinical diagnostic microbiology," *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, vol. 1854, no. 6, pp. 528–537, 2015.

[4] T. Brignoli, M. Recker, W. W. Lee, T. Dong, R. Bhamber, M. Albur, P. Williams, A. W. Dowsey, and R. C. Massey, "Diagnostic maldi-tof ms can differentiate between high and low toxic staphylococcus aureus bacteraemia isolates as a predictor of patient outcome," *Microbiology*, vol. 168, no. 8, p. 001223, 2022.

[5] N. Dhiman, L. Hall, S. L. Wohlfiel, S. P. Buckwalter, and N. L. Wengenack, "Performance and cost analysis of matrix-assisted laser desorption ionization–time of flight mass spectrometry for routine identification of yeast," *Journal of clinical microbiology*, vol. 49, no. 4, pp. 1614–1616, 2011.

[6] R. Holland, J. Wilkes, F. Rafii, J. Sutherland, C. Persons, K. Voorhees, and J. O. Lay, Jr., "Rapid identification of intact whole bacteria based on spectral patterns using matrix-assisted laser desorption/ionization with time-of-flight mass spectrometry," *Rapid Communications in Mass Spectrometry*, vol. 10, no. 10, pp. 1227–1232, 1996.

[7] C. Fenselau and P. A. Demirev, "Characterization of intact microorganisms by maldi mass spectrometry," *Mass Spectrometry Reviews*, vol. 20, no. 4, pp. 157–171, 2001.

[8] R. Giebel, C. Worden, S. Rust, G. Kleinheinz, M. Robbins, and T. R. Sandrin, "Microbial fingerprinting using matrix-assisted laser desorption ionization time-of-flight mass spectrometry (maldi-tof ms): applications and challenges," *Advances in Applied Microbiology*, vol. 71, pp. 149–184, 2010.

[9] W. H. Organization, "Antimicrobial resistance," Tech. Rep., 2023. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/antimicrobial-resistance

[10] L. B. Reller, M. Weinstein, J. H. Jorgensen, and M. J. Ferraro, "Antimicrobial susceptibility testing: a review of general principles and contemporary practices," *Clinical infectious diseases*, vol. 49, no. 11, pp. 1749–1755, 2009.

[11] C. V. Weis, C. R. Jutzeler, and K. Borgwardt, "Machine learning for microbial identification and antimicrobial susceptibility testing on maldi-tof mass spectra: a systematic review," *Clinical Microbiology and Infection*, vol. 26, no. 10, pp. 1310–1317, 2020.

[12] A. Van Belkum, M. Struelens, A. De Visser, H. Verbrugh, and M. Tibayrenc, "Guidelines for the validation and application of typing methods for use in bacterial epidemiology," *Clinical Microbiology and Infection*, vol. 13, no. s3, pp. 1–46, 2007. [Online]. Available: https://doi.org/10.1111/j.1469-0691.2007.01786.x

[13] A. Roetzer, R. Diel, T. A. Kohl, C. Rückert, U. Nübel, J. Blom, T. Wirth, S. Jaenicke, S. Schuback, S. Rüsch-Gerdes *et al.*, "Whole genome sequencing versus traditional genotyping for investigation of a mycobacterium tuberculosis outbreak: a longitudinal molecular epidemiological study," *PLoS medicine*, vol. 10, no. 2, p. e1001387, 2013.

[14] World Health Organization, "Antimicrobial resistance, hypervirulent klebsiella pneumoniae - global situation," 2024, accessed: 2024-12-03. [Online]. Available: https://www.who.int/emergencies/disease-outbreak-news/item/2024-DON527

[15] H. Desaire and D. Hua, "The aristotle classifier provides accurate identifications of highly similar bacteria analyzed by maldi-tof ms," *Analytical Chemistry*, 11 2020.

[16] T. Mortier, A. Wieme, P. Vandamme, and W. Waegeman, "Bacterial species identification using maldi-tof mass spectrometry and machine learning techniques: A large-scale benchmarking study," *Computational and Structural Biotechnology Journal*, vol. 19, 11 2021.

[17] Y. Li, Z. Gan, X. Zhou, and Z. Chen, "Accurate classification of listeria species by maldi-tof mass spectrometry incorporating denoising autoencoder and machine learning," *Journal of Microbiological Methods*, vol. 192, 2022.

[18] A. Candela, A. Guerrero-López, M. Mateos, A. Gómez-Asenjo, M. J. Arroyo, M. Hernandez-García, R. del Campo, E. Cercenado, A. Cuénod, G. Méndez, L. Mancera, J. de Dios Caballero, L. Martínez-García, D. Gijón, M. I. Morosini, P. Ruiz-Garbajosa, A. Egli, R. Cantón, P. Muñoz, D. Rodríguez-Temporal, and B. Rodríguez-Sánchez, "Automatic discrimination of species within the enterobacter cloacae complex using matrix-assisted laser desorption ionization–time of flight mass spectrometry and supervised algorithms," *Journal of Clinical Microbiology*, vol. 61, no. 4, 2023.

[19] Q. Mao, X. Zhang, Z. Xu, Y. Xiao, Y. Song, and F. Xu, "Identification of escherichia coli strains using maldi-tof ms combined with long short-term memory neural networks," *Aging*, vol. 16, 06 2024.

[20] T.-S. Huang, S. S.-J. Lee, C.-C. Lee, and F.-C. Chang, "Detection of carbapenem-resistant klebsiella pneumoniae on the basis of matrix-assisted laser desorption ionization time-of-flight mass spectrometry by using supervised machine learning approach," *PLOS ONE*, vol. 15, 02 2020.

[21] A. Calderaro, M. Buttrini, B. Farina, S. Montecchini, M. Martinelli, F. Crocamo, M. Arcangeletti, C. Chezzi, and F. De Conto, "Rapid identification of escherichia coli colistin-resistant strains by maldi-tof mass spectrometry," *Microorganisms*, vol. 9, 10 2021.

[22] C. Weis, A. Cuénod, B. Rieck, O. Dubuis, S. Graf, C. Lang, M. Oberle, M. Brackmann, K. Søgaard, M. Osthoff, K. Borgwardt, and A. Egli, "Direct antimicrobial resistance prediction from clinical maldi-tof mass spectra using machine learning," *Nature Medicine*, vol. 28, 01 2022.

[23] Y. Zeng, C. Wang, Q. Ye, G. Liu, L. Zhang, J. Wan, and Y. Zhu, "Machine learning model of imipenem-resistant klebsiella pneumoniae based on maldi-tof-ms platform: An observational study," *Health Science Reports*, vol. 6, no. 9, 2023.

[24] T.-H. Lin, H.-Y. Chung, M.-J. Jian, C.-K. Chang, H.-H. Lin, C.-M. Yu, C.-L. Perng, F.-Y. Chang, C.-W. Chen, and H.-S. Shang, "Innovative strategies against superbugs: Developing an ai-cdss for precise stenotrophomonas maltophilia treatment," *Journal of Global Antimicrobial Resistance*, vol. 38, pp. 173–180, 2024.

[25] Y. Huang and H. Desaire, "Calibration-free quantification of antibiotic resistance markers using machine learning on maldi-tof data," *Analytical Chemistry*, vol. 92, no. 2, pp. 1050–1057, 2020.

[26] C. Weis, M. Horn, B. Rieck, A. Cuénod, A. Egli, and K. Borgwardt, "Topological and kernel-based microbial phenotype prediction from MALDI-TOF mass spectra," *Bioinformatics*, vol. 36, no. Supplement1, 07 2020.

[27] L. Qiao, "Maldi-tof ms for pathogenic bacteria analysis," *International Journal of Mass Spectrometry*, vol. 482, p. 116935, 2022.

[28] I. Pena, E. Pena-Vina, I. Rodríguez-Avial, J. Picazo, á. Gómez-González, and E. Culebras, "Comparison of performance of maldi-tof ms and mlst for biotyping carbapenemase-producing klebsiella pneumoniae sequence types st11 and st101 isolates," *Enfermedades Infecciosas y Microbiología Clínica*, vol. 40, 12 2020.

[29] A. Calderaro, M. Buttrini, M. Martinelli, B. Farina, T. Moro, S. Montecchini, M. Arcangeletti, C. Chezzi, and F. De Conto, "Rapid classification of clostridioides difficile strains using maldi-tof ms peak-based assay in comparison with pcr-ribotyping," *Microorganisms*, vol. 9, 03 2021.

[30] A. Abdrabou, I. Sy, M. Bischoff, M. Arroyo, B. L. Sören, A. Mellmann, L. von Müller, B. Gärtner, and F. Berger, "Discrimination between hypervirulent and non-hypervirulent ribotypes of clostridioides difficile by maldi-tof mass spectrometry and machine learning," *European Journal of Clinical Microbiology & Infectious Diseases*, 09 2023.

[31] E. Manfredi, F. Rocca, J. Zintgraff, L. Irazu, E. Miliwebsky, C. Carbonari, N. Deza, M. Prieto, and I. Chinen, "Rapid and accurate detection of shiga toxin-producing escherichia coli (stec) serotype o157 : H7 by mass spectrometry directly from the isolate, using 10 potential biomarker peaks and machine learning predictive models," *Journal of medical microbiology*, vol. 72, 05 2023.

[32] A. Godmer, Q. Giai Gianetto, K. Le Neindre, V. Latapy, M. Bastide, M. Ehmig, V. Lalande, N. Veziris, A. Aubry, F. Barbut, and C. Eckert, "Contribution of maldi-tof mass spectrometry and machine learning including deep learning techniques for the detection of virulence factors of clostridioides difficile strains," *Microbial Biotechnology*, vol. 17, no. 6, 2024.

[33] C. Weis, A. Cuénod, B. Rieck *et al.*, "DRIAMS: Database of Resistance Information on Antimicrobials and MALDI-TOF Mass Spectra," https://doi.org/10.5061/dryad.bzkh1899q, 2022.

[34] P. Lasch, M. Stämmler, and A. Schneider, "A MALDI-TOF Mass Spectrometry Database for Identification and Classification of Highly Pathogenic Microorganisms from the Robert Koch-Institute (RKI)," Oct. 2016. [Online]. Available: https://doi.org/10.5281/zenodo.163517

[35] X. A. López-Cortés, J. M. Manríquez-Troncoso, J. Kandalaft-Letelier, and S. Cuadros-Orellana, "Machine learning and matrix-assisted laser desorption/ionization time-of-flight mass spectra for antimicrobial resistance prediction: A systematic review of recent advancements and future development," *Journal of Chromatography A*, p. 465262, 2024.

[36] F. Pennisi, A. Pinto, G. E. Ricciardi, C. Signorelli, and V. Gianfredi, "Artificial intelligence in antimicrobial stewardship: a systematic review and meta-analysis of predictive performance and diagnostic accuracy," *European Journal of Clinical Microbiology & Infectious Diseases*, pp. 1–51, 2025.

[37] Z. Zhang, D. Wang, P. de B Harrington, K. J. Voorhees, and J. Rees, "Forward selection radial basis function networks applied to bacterial classification based on maldi-tof-ms," *Talanta*, vol. 63, no. 3, pp. 527–532, 2004.

[38] G. A. Satten, S. Datta, H. Moura, A. R. Woolfitt, M. d. G. Carvalho, G. M. Carlone, B. K. De, A. Pavlopoulos, and J. R. Barr, "Standardization and denoising algorithms for mass spectra to classify whole-organism bacterial specimens," *Bioinformatics*, vol. 20, no. 17, pp. 3128–3136, 2004.

[39] S.-Y. Hsieh, C.-L. Tseng, Y.-S. Lee, A.-J. Kuo, C.-F. Sun, Y.-H. Lin, and J.-K. Chen, "Highly efficient classification and identification of human pathogenic bacteria by maldi-tof ms," *Molecular & cellular proteomics*, vol. 7, no. 2, pp. 448–456, 2008.

[40] R. Schaumann, N. Knoop, G. H. Genzel, K. Losensky, C. Rosenkranz, C. S. Stîngu, W. Schellenberger, A. C. Rodloff, and K. Eschrich, "A step towards the discrimination of beta-lactamase-producing clinical isolates of enterobacteriaceae and pseudomonas aeruginosa by maldi-tof mass spectrometry," *Medical science monitor: international medical journal of experimental and clinical research*, vol. 18, no. 9, p. MT71, 2012.

[41] P. D. Khot and M. A. Fisher, "Novel approach for differentiating shigella species and escherichia coli by matrix-assisted laser desorption ionization–time of flight mass spectrometry," *Journal of clinical microbiology*, vol. 51, no. 11, pp. 3711–3716, 2013.

[42] S. Nakano, Y. Matsumura, K. Kato, T. Yunoki, G. Hotta, T. Noguchi, M. Yamamoto, M. Nagao, Y. Ito, S. Takakura *et al.*, "Differentiation of vana-positive enterococcus faecium from vana-negative e. faecium by matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry," *International journal of antimicrobial agents*, vol. 44, no. 3, pp. 256–259, 2014.

[43] B. Li, T. Guo, F. Qu, B. Li, H. Wang, Z. Sun, X. Li, Z. Gao, C. Bao, C. Zhang *et al.*, "Matrix-assisted laser desorption ionization: time of flight mass spectrometry-identified models for detection of esbl-producing bacterial strains," *Medical Science Monitor Basic Research*, vol. 20, p. 176, 2014.

[44] D. Xiao, F. Zhao, H. Zhang, F. Meng, and J. Zhang, "Novel strategy for typing mycoplasma pneumoniae isolates by use of matrix-assisted laser desorption ionization–time of flight mass spectrometry coupled with clinprotools," *Journal of clinical microbiology*, vol. 52, no. 8, pp. 3038–3043, 2014.

[45] T. Zhang, J. Ding, X. Rao, J. Yu, M. Chu, W. Ren, L. Wang, and W. Xue, "Analysis of methicillin-resistant staphylococcus aureus major clonal lineages by matrix-assisted laser desorption ionization–time of flight mass spectrometry (maldi–tof ms)," *Journal of microbiological methods*, vol. 117, pp. 122–127, 2015.

[46] S. Nakano, Y. Matsumura, Y. Ito, T. Fujisawa, B. Chang, S. Suga, K. Kato, T. Yunoki, G. Hotta, T. Noguchi *et al.*, "Development and evaluation of maldi-tof ms-based serotyping for streptococcus pneumoniae," *European Journal of Clinical Microbiology & Infectious Diseases*, vol. 34, pp. 2191–2198, 2015.

[47] J. H. Chen, K. K. She, O.-Y. Wong, J. L. Teng, W.-C. Yam, S. K. Lau, P. C. Woo, V. C. Cheng, and K.-Y. Yuen, "Use of maldi biotyper

plus clinprotools mass spectra analysis for correct identification of streptococcus pneumoniae and streptococcus mitis/oralis," *Journal of Clinical Pathology*, vol. 68, no. 8, pp. 652–656, 2015.

[48] M. Camoez, J. Sierra, M. Dominguez, M. Ferrer-Navarro, J. Vila, and I. Roca, "Automated categorization of methicillin-resistant staphylococcus aureus clinical isolates into different clonal complexes by maldi-tof mass spectrometry," *Clinical Microbiology and Infection*, vol. 22, no. 2, pp. 161–e1, 2016.

[49] C. A. Mather, B. J. Werth, S. Sivagnanam, D. J. SenGupta, and S. M. Butler-Wu, "Rapid detection of vancomycin-intermediate staphylococcus aureus by matrix-assisted laser desorption ionization–time of flight mass spectrometry," *Journal of clinical microbiology*, vol. 54, no. 4, pp. 883–890, 2016.

[50] P.-L. Ho, C.-Y. Yau, L.-Y. Ho, J. H. Chen, E. L. Lai, S. W. Lo, W. Cindy, and K.-H. Chow, "Rapid detection of cfia metallo-β-lactamase-producing bacteroides fragilis by the combination of maldi-tof ms and carbanp," *Journal of Clinical Pathology*, vol. 70, no. 10, pp. 868–873, 2017.

[51] K. Sogoawa, M. Watanabe, T. Ishige, S. Segawa, A. Miyabe, S. Murata *et al.*, "Rapid staphylococcus aureus discrimination between methicillin-sensitive and methicillin-resistant using maldi-tof mass spectrometry," *Biocontrol Sci*, vol. 22, no. 163, p. e9, 2017.

[52] H.-Y. Wang, T.-Y. Lee, Y.-J. Tseng, T.-P. Liu, K.-Y. Huang, Y.-T. Chang, C.-H. Chen, and J.-J. Lu, "A new scheme for strain typing of methicillin-resistant staphylococcus aureus on the basis of matrix-assisted laser desorption ionization time-of-flight mass spectrometry by using machine learning approach," *PloS one*, vol. 13, no. 3, p. e0194289, 2018.

[53] H.-Y. Wang, F. Lien, T.-P. Liu, C.-H. Chen, C.-J. Chen, and J.-J. Lu, "Application of a maldi-tof analysis platform (clinprotools) for rapid and preliminary report of mrsa sequence types in taiwan," *PeerJ*, vol. 6, p. e5784, 2018.

[54] C. Rodrigues, V. Passet, A. Rakotondrasoa, and S. Brisse, "Identification of klebsiella pneumoniae, klebsiella quasipneumoniae, klebsiella variicola and related phylogroups by maldi-tof mass spectrometry," *Frontiers in microbiology*, vol. 9, p. 3000, 2018.

[55] K. Asakura, T. Azechi, H. Sasano, H. Matsui, H. Hanaki, M. Miyazaki, T. Takata, M. Sekine, T. Takaku, T. Ochiai *et al.*, "Rapid and easy detection of low-level resistance to vancomycin in methicillin-resistant staphylococcus aureus by matrix-assisted laser desorption ionization time-of-flight mass spectrometry," *PloS one*, vol. 13, no. 3, p. e0194212, 2018.

[56] H.-Y. Wang, C.-H. Chen, T.-Y. Lee, J.-T. Horng, T.-P. Liu, Y.-J. Tseng, and J.-J. Lu, "Rapid detection of heterogeneous vancomycin-intermediate staphylococcus aureus based on matrix-assisted laser desorption ionization time-of-flight: using a machine learning approach and unbiased validation," *Frontiers in Microbiology*, vol. 9, p. 2393, 2018.

[57] J. Ling, H. Wang, G. Li, Z. Feng, Y. Song, P. Wang, H. Shao, H. Zhou, and G. Chen, "A novel short-term high-lactose culture approach combined with a matrix-assisted laser desorption ionization-time of flight mass spectrometry assay for differentiating escherichia coli and shigella species using artificial neural networks," *PLoS One*, vol. 14, no. 10, p. e0222636, 2019.

[58] C.-R. Chung, H.-Y. Wang, F. Lien, Y.-J. Tseng, C.-H. Chen, T.-Y. Lee, T.-P. Liu, J.-T. Horng, and J.-J. Lu, "Incorporating statistical test and machine intelligence into strain typing of staphylococcus haemolyticus based on matrix-assisted laser desorption ionization-time of flight mass spectrometry," *Frontiers in microbiology*, vol. 10, p. 2120, 2019.

[59] W. Tang, N. Ranganathan, V. Shahrezaei, and G. Larrouy-Maumus, "Maldi-tof mass spectrometry on intact bacteria combined with a refined analysis framework allows accurate classification of mssa and mrsa," *PloS one*, vol. 14, no. 6, p. e0218951, 2019.

[60] H.-Y. Wang, W.-C. Li, K.-Y. Huang, C.-R. Chung, J.-T. Horng, J.-F. Hsu, J.-J. Lu, and T.-Y. Lee, "Rapid classification of group b streptococcus serotypes based on matrix-assisted laser desorption ionization-time of flight mass spectrometry and machine learning techniques," *BMC bioinformatics*, vol. 20, pp. 1–17, 2019.

[61] J. Ling, G. Li, H. Shao, H. Wang, H. Yin, H. Zhou, Y. Song, and G. Chen, "Helix matrix transformation combined with convolutional neural network algorithm for matrix-assisted laser desorption ionization-time of flight mass spectrometry-based bacterial identification," *Frontiers in Microbiology*, vol. 11, 11 2020.

[62] S. Kann, S. Sao, C. Phoeung, Y. By, J. Bryant, F. Komurian-Pradel, V. Saphonn, M. Chou, and P. Turner, "Maldi-tof mass spectrometry for sub-typing of streptococcus pneumoniae," *BMC microbiology*, vol. 20, pp. 1–9, 2020.

[63] J. Wei, H. Zhang, H. Zhang, E. Zhang, B. Zhang, F. Zhao, and D. Xiao, "Novel strategy for rapidly and safely distinguishing bacillus anthracis and bacillus cereus by use of peptide mass fingerprints based on matrix-assisted laser desorption ionization–time of flight mass spectrometry," *Journal of Clinical Microbiology*, vol. 59, no. 1, 2020.

[64] H.-Y. Wang, C.-R. Chung, Z. Wang, S. Li, B.-Y. Chu, J.-T. Horng, J.-J. Lu, and T.-Y. Lee, "A large-scale investigation and identification of methicillin-resistant Staphylococcus aureus based on peaks binning of matrix-assisted laser desorption ionization-time of flight MS spectra," *Briefings in Bioinformatics*, vol. 22, no. 3, 07 2021.

[65] E. Gato, I. P. Constanso, A. Candela, F. Galán, B. K. Rodiño-Janeiro, M. J. Arroyo, G. Méndez, L. Mancera, T. Alioto, M. Gut, I. Gut, M. álvarez Tejado, B. Rodríguez-Sánchez, G. Bou, and M. Oviaño, "An improved matrix-assisted laser desorption ionization–time of flight mass spectrometry data analysis pipeline for the identification of carbapenemase-producing klebsiella pneumoniae," *Journal of Clinical Microbiology*, vol. 59, no. 7, 2021.

[66] H.-Y. Wang, C.-R. Chung, C.-J. Chen, K.-P. Lu, Y.-J. Tseng, T.-H. Chang, M.-H. Wu, W.-T. Huang, T.-W. Lin, T.-P. Liu, T.-Y. Lee, J.-T. Horng, and J.-J. Lu, "Clinically applicable system for rapidly predicting enterococcus faecium susceptibility to vancomycin," *Microbiology Spectrum*, vol. 9, no. 3, 2021.

[67] S. Fiamanya, L. Cipolla, M. Prieto, and J. Stelling, "Exploring the value of maldi-tof ms for the detection of clonal outbreaks of burkholderia contaminans," *Journal of microbiological methods*, vol. 181, p. 106130, 2021.

[68] K. Tang, D. Tang, Q. Wang, and C. Li, "Maldi-tof ms platform combined with machine learning to establish a model for rapid identification of methicillin-resistant staphylococcus aureus," *Journal of Microbiological Methods*, vol. 180, 2021.

[69] C.-R. Chung, Z. Wang, J.-M. Weng, H.-Y. Wang, L.-C. Wu, Y.-J. Tseng, C.-H. Chen, J.-J. Lu, J.-T. Horng, and T.-Y. Lee, "Mdrsa: A web based-tool for rapid identification of multidrug resistant staphylococcus aureus based on matrix-assisted laser desorption ionization-time of flight mass spectrometry," *Frontiers in Microbiology*, vol. 12, 12 2021.

[70] L. Huang, K. Gao, G. Chen, H. Zhong, Z. Li, X. Guan, Q. Deng, Y. Xie, W. Ji, D. McIver, C.-Y. Chang, and H. Liu, "Rapid classification of multilocus sequence subtype for group b streptococcus based on maldi-tof mass spectrometry and statistical models," *Frontiers in Cellular and Infection Microbiology*, vol. 10, 01 2021.

[71] X. Liu, T. Su, Y.-M. S. Hsu, H. Yu, H. S. Yang, L. Jiang, and Z. Zhao, "Rapid identification and discrimination of methicillin-resistant strains via matrix-assisted laser desorption/ionization time-of-flight mass spectrometry," *Rapid Communications in Mass Spectrometry*, vol. 35, no. 2, 2021.

[72] M. Feucherolles, H.-M. Cauchie, M. Nennig, S. L. Becker, D. Martiny, S. Losch, C. Penny, and C. Ragimbeau, "Combination of maldi-tof mass spectrometry and machine learning for rapid antimicrobial resistance screening: The case of campylobacter spp," *Frontiers in Microbiology*, vol. 12, 02 2022.

[73] P.-H. Kong, C.-H. Chiang, T.-C. Lin, S.-C. Kuo, C.-F. Li, C. Hsiung, Y.-L. Shiue, Y. Yao, L.-C. Wu, and H.-H. Tsou, "Discrimination of methicillin-resistant staphylococcus aureus by maldi-tof mass spectrometry with machine learning techniques in patients with staphylococcus aureus bacteremia," *Pathogens*, vol. 11, 05 2022.

[74] H.-Y. Wang, T. T. Hsieh, C.-R. Chung, H.-C. Chang, J.-T. Horng, J.-J. Lu, and J.-H. Huang, "Efficiently predicting vancomycin resistance of enterococcus faecium from maldi-tof ms spectra using a deep learning-based approach," *Frontiers in Microbiology*, vol. 13, 06 2022.

[75] C. Wang, Z. Wang, H.-Y. Wang, C.-R. Chung, J.-T. Horng, J.-J. Lu, and T.-Y. Lee, "Large-scale samples based rapid detection of ciprofloxacin resistance in klebsiella pneumoniae using machine learning methods," *Frontiers in Microbiology*, vol. 13, 03 2022.

[76] F. Dematheis, M. Walter, D. Lang, M. Antwerpen, H. Scholz, M.-T. Pfalzgraf, E. Mantel, C. Hinz, R. Wölfel, and S. Zange, "Machine learning algorithms for classification of maldi-tof ms spectra from phylogenetically closely related species brucella melitensis, brucella abortus and brucella suis," *Microorganisms*, vol. 10, 08 2022.

[77] M. J. van den Beld, J. W. Rossen, N. Evers, M. A. Kooistra-Smid, and F. A. Reubsaet, "Maldi-tof ms using a custom-made database, biomarker assignment, or mathematical classifiers does not differentiate shigella spp. and escherichia coli," *Microorganisms*, vol. 10, no. 2, p. 435, 2022.

[78] K. Jeon, J.-M. Kim, K. Rho, S. Jung, H. Park, and J.-S. Kim, "Performance of a machine learning-based methicillin resistance of staphylococcus aureus identification system using maldi-tof ms and

comparison of the accuracy according to sccmec types," *Microorganisms*, vol. 10, 09 2022.

[79] A. Candela, M. Arroyo, á. Sánchez-Molleda, G. Méndez Cervantes, L. Quiroga, A. Ruiz, E. Cercenado, M. Marin, P. Muñoz, L. Mancera, D. Rodriguez-Temporal, and B. Rodriguez-Sanchez, "Rapid and reproducible maldi-tof-based method for the detection of vancomycin-resistant enterococcus faecium using classifying algorithms," *Diagnostics*, vol. 12, 01 2022.

[80] J. Zhang, Z. Wang, H.-Y. Wang, C.-R. Chung, J.-T. Horng, J.-J. Lu, and T.-Y. Lee, "Rapid antibiotic resistance serial prediction in staphylococcus aureus based on large-scale maldi-tof data by applying xgboost in multi-label learning," *Frontiers in Microbiology*, vol. 13, 04 2022.

[81] J. Wang, C. Xia, Y. Wu, X. Tian, K. Zhang, and Z. Wang, "Rapid detection of carbapenem-resistant klebsiella pneumoniae using machine learning and maldi-tof ms platform," *Infection and Drug Resistance*, vol. Volume 15, 07 2022.

[82] J. Yu, N. Tien, Y.-C. Liu, D.-Y. Cho, J.-W. Chen, Y.-T. Tsai, Y.-C. Huang, H.-J. Chao, and C.-J. Chen, "Rapid identification of methicillin-resistant staphylococcus aureus using maldi-tof ms and machine learning from over 20,000 clinical isolates," *Microbiology Spectrum*, vol. 10, 03 2022.

[83] á. Rodríguez-Villodres, L. Galvez-Benitez, M. Arroyo, G. Méndez Cervantes, L. Mancera, A. Vila, J. Jímenez, and Y. Smani, "Ultrasensitive and rapid identification of esri developer- and piperacillin/tazobactam-resistant escherichia coli by the maldipiptaz test," *Emerging Microbes & Infections*, vol. 11, 08 2022.

[84] Z. Wang, Y. Pang, C.-R. Chung, H.-Y. Wang, H. Cui, Y.-C. Chiang, J.-T. Horng, J.-J. Lu, and T.-Y. Lee, "A risk assessment framework for multidrug-resistant Staphylococcus aureus using machine learning and mass spectrometry technology," *Briefings in Bioinformatics*, vol. 24, no. 6, 09 2023.

[85] A. Guerrero-López, C. Sevilla-Salcedo, A. Candela, M. Hernández-García, E. Cercenado, P. M. Olmos, R. Cantón, P. Muñoz, V. Gómez-Verdejo, R. del Campo *et al.*, "Automatic antibiotic resistance prediction in klebsiella pneumoniae based on maldi-tof mass spectra," *Engineering Applications of Artificial Intelligence*, vol. 118, p. 105644, 2023.

[86] C.-R. Chung, H.-Y. Wang, C.-H. Yao, L.-C. Wu, J.-J. Lu, J.-T. Horng, and T.-Y. Lee, "Data-driven two-stage framework for identification and characterization of different antibiotic-resistant escherichia coli isolates based on mass spectrometry data," *Microbiology spectrum*, vol. 11, 04 2023.

[87] E. Gato, M. Arroyo, G. Méndez Cervantes, A. Candela, B. Rodiño-Janeiro, J. Fernández, B. Rodriguez-Sanchez, L. Mancera, J. Arca-Suárez, A. Beceiro, G. Bou, and M. Oviaño, "Direct detection of carbapenemase-producing klebsiella pneumoniae by maldi-tof analysis of full spectra applying machine learning," *Journal of clinical microbiology*, vol. 61, 05 2023.

[88] J. Yu, Y.-T. Lin, W.-C. Chen, K.-H. Tseng, H.-H. Lin, N. Tien, C.-F. Cho, J.-Y. Huang, S.-J. Liang, L.-C. Ho, Y.-W. Hsieh, K.-C. Hsu, M.-W. Ho, P.-R. Hsueh, and D.-Y. Cho, "Direct prediction of carbapenem-resistant, carbapenemase-producing, and colistin-resistant klebsiella pneumoniae isolates from routine maldi-tof mass spectra using machine learning and outcome evaluation," *International Journal of Antimicrobial Agents*, vol. 61, no. 6, 2023.

[89] J. Zintgraff, F. Rocca, N. S. Eluchans, L. Irazu, M. A. Moscoloni, C. Lara, and M. Santos, "Exploring streptococcus pneumoniae capsular typing through maldi-tof mass spectrometry and machine-learning algorithms in argentina: Identifying prevalent non pcv13 serotypes alongside pcv13 serotypes," *Journal of Mass Spectrometry and Advances in the Clinical Lab*, vol. 30, 2023.

[90] D. Rodríguez-Temporal, L. Herrera, F. Alcaide, D. Domingo, G. Héry-Arnaud, J. van Ingen, A. V. den Bossche, A. Ingebretsen, C. Beauruelle, E. Terschlüsen, S. Boarbi, N. Vila, M. J. Arroyo, G. Méndez, P. Muñoz, L. Mancera, M. J. Ruiz-Serrano, and B. Rodríguez-Sánchez, "Identification of mycobacterium abscessus subspecies by maldi-tof mass spectrometry and machine learning," *Journal of Clinical Microbiology*, vol. 61, no. 1, 2023.

[91] J. Yu, H.-H. Lin, K.-H. Tseng, Y.-T. Lin, W.-C. Chen, N. Tien, C.-F. Cho, S.-J. Liang, L.-C. Ho, Y.-W. Hsieh, K.-C. Hsu, M.-W. Ho, P.-R. Hsueh, and D.-Y. Cho, "Prediction of methicillin-resistant staphylococcus aureus and carbapenem-resistant klebsiella pneumoniae from flagged blood cultures by combining rapid sepsityper maldi-tof mass spectra with machine learning," *International Journal of Antimicrobial Agents*, vol. 62, 10 2023.

[92] Z. Wang, Y. Zhou, G. Guo, Q. Li, Y. Yu, and W. Zhang, "Promising potential of machine learning-assisted maldi-tof ms as an effective detector for ¡i¿streptococcus suis¡/i¿ serotype 2 and virulence thereof," *Applied and Environmental Microbiology*, vol. 89, no. 11, 2023.

[93] J. Fu, F. He, J. Xiao, Z. Liao, L. He, J. He, J. Guo, and S. Liu, "Rapid amr prediction in pseudomonas aeruginosa combining maldi-tof ms with dnn model," *Journal of applied microbiology*, vol. 134, 11 2023.

[94] H.-Y. Wang, C.-H. Kuo, C.-R. Chung, W.-Y. Lin, Y.-C. Wang, T.-W. Lin, J.-R. Yu, J.-J. Lu, and T.-S. Wu, "Rapid and accurate discrimination of mycobacterium abscessus subspecies based on matrix-assisted laser desorption ionization-time of flight spectrum and machine learning algorithms," *Biomedicines*, vol. 11, no. 1, p. 45, 2023.

[95] S. Iskender, S. Heydarov, M. Yalcin, C. Faydaci, O. Kurt, S. Surme, and O. Kucukbasmaci, "Rapid determination of colistin resistance in klebsiella pneumoniae by maldi-tof peak based machine learning algorithm with matlab," *Diagnostic Microbiology and Infectious Disease*, 08 2023.

[96] K. Liu, Y. Wang, M. Zhao, G. Xue, A. Wang, W. Wang, L. Xu, and J. Chen, "Rapid discrimination of bifidobacterium longum subspecies based on maldi-tof ms and machine learning," *Frontiers in Microbiology*, vol. 14, 12 2023.

[97] A. Candela, M. J. Arroyo, M. Sánchez-Cueto, M. Marín, E. Cercenado, G. Méndez, P. Muñoz, L. Mancera, D. Rodríguez-Temporal, and B. Rodríguez-Sánchez, "Rapid discrimination of pseudomonas aeruginosa st175 isolates involved in a nosocomial outbreak using maldi-tof mass spectrometry and ftir spectroscopy coupled with machine learning," *Transboundary and Emerging Diseases*, vol. 2023, no. 1, p. 8649429, 2023.

[98] Y.-M. Zhang, M.-F. Tsao, C.-Y. Chang, K.-T. Lin, J. Keller, and H.-C. Lin, "Rapid identification of carbapenem-resistant klebsiella pneumoniae based on matrix-assisted laser desorption ionization time-of-flight mass spectrometry and an artificial neural network model," *Journal of Biomedical Science*, vol. 30, 04 2023.

[99] C.-R. Chung, H.-Y. Wang, P.-H. Chou, L.-C. Wu, J.-J. Lu, J.-T. Horng, and T.-Y. Lee, "Towards accurate identification of antibiotic-resistant pathogens through the ensemble of multiple preprocessing methods based on maldi-tof spectra," *International Journal of Molecular Sciences*, vol. 24, 01 2023.

[100] W.-Y. Wang, C.-F. Chiu, S.-M. Tsao, Y.-L. Lee, and Y.-H. Chen, "Accurate prediction of antimicrobial resistance and genetic marker of staphylococcus aureus clinical isolates using maldi-tof ms and machine learning–across driams and taiwan database," *International Journal of Antimicrobial Agents*, vol. 64, no. 5, p. 107329, 2024.

[101] G. De Waele, G. Menschaert, and W. Waegeman, "An antimicrobial drug recommender system using maldi-tof ms and dual-branch neural networks," *eLife*, vol. 13, p. RP93242, 2024.

[102] A. Candela, D. Rodriguez-Temporal, M. Blázquez-Sánchez, M. J. Arroyo, M. Marín, L. Alcalá, G. Bou, B. Rodríguez-Sánchez, and M. Oviaño, "Analysis of high-molecular-weight proteins using maldi-tof ms and machine learning for the differentiation of clinically relevant clostridioides difficile ribotypes," *European Journal of Clinical Microbiology & Infectious Diseases*, pp. 1–9, 2024.

[103] T.-H. Lin, H.-Y. Chung, M.-J. Jian, C.-K. Chang, H.-H. Lin, C.-M. Yu, C.-L. Perng, F.-Y. Chang, C.-W. Chen, C.-H. Chiu *et al.*, "Artificial intelligence-clinical decision support system for enhanced infectious disease management: Accelerating ceftazidime-avibactam resistance detection in klebsiella pneumoniae," *Journal of Infection and Public Health*, vol. 17, no. 10, p. 102541, 2024.

[104] M.-J. Jian, T.-H. Lin, H.-Y. Chung, C.-K. Chang, C.-L. Perng, F.-Y. Chang, and H.-S. Shang, "Artificial intelligence-clinical decision support system in infectious disease control: combatting multidrug-resistant klebsiella pneumoniae with machine learning," *Infection and Drug Resistance*, pp. 2899–2912, 2024.

[105] L. Huang, K. Gao, H. Zhong, Y. Xie, B. Liang, W. Ji, and H. Liu, "Automated classification of group b streptococcus into different clonal complexes using maldi-tof mass spectrometry," *Frontiers in Molecular Biosciences*, vol. 11, p. 1355448, 2024.

[106] X. Xu, "Modelling the rapid detection of carbapenemase-resistant klebsiella pneumoniae based on machine learning and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry," *Diagnostic Microbiology and Infectious Disease*, vol. 110, no. 2, p. 116467, 2024.

[107] X. López, J. Manriquez, R. Hernández García, and D. Peralta, "Ms-deepamr: antimicrobial resistance prediction based on deep neural networks and transfer learning," *Frontiers in Microbiology*, vol. 15, 04 2024.

[108] C. A. Astudillo, X. A. López-Cortés, E. Ocque, and J. M. Manríquez-Troncoso, "Multi-label classification to predict antibiotic resistance from raw clinical maldi-tof mass spectrometry data," *Scientific Reports*, vol. 14, no. 1, p. 31283, 2024.

[109] M.-J. Jian, T.-H. Lin, H.-Y. Chung, C.-K. Chang, C.-L. Perng, F.-Y. Chang, and H.-S. Shang, "Pioneering klebsiella pneumoniae antibiotic resistance prediction with artificial intelligence-clinical decision support system–enhanced matrix-assisted laser desorption/ionization time-of-flight mass spectrometry: Retrospective study," *Journal of Medical Internet Research*, vol. 26, p. e58039, 2024.

[110] V. Macaya Mejias, D. Zabala-Blanco, X. A. López-Cortés, F. Tirado, J. M. Manriquez-Troncoso, and R. Ahumada-Garcia, "Predicting bacterial antibiotic resistance using maldi-tof mass spectrometry databases with elm applications," *Journal of Computer Science & Technology*, vol. 24, 2024.

[111] H.-A. Nguyen, A. Y. Peleg, J. Song, B. Antony, G. I. Webb, J. A. Wisniewski, L. V. Blakeway, G. Z. Badoordeen, R. Theegala, H. Zisis *et al.*, "Predicting pseudomonas aeruginosa drug resistance using artificial intelligence and clinical maldi-tof mass spectra," *MSystems*, vol. 9, no. 9, pp. e00 789–24, 2024.

[112] M. Ren, Q. Chen, and J. Zhang, "Repurposing maldi-tof ms for effective antibiotic resistance screening in staphylococcus epidermidis using machine learning," *Scientific Reports*, vol. 14, no. 1, p. 24139, 2024.

[113] S. Gibb and K. Strimmer, "Maldiquant: a versatile r package for the analysis of mass spectrometry data," *Bioinformatics*, vol. 28, no. 17, pp. 2270–2271, 2012.

[114] R. Schaumann, K. Dallacker-Losensky, C. Rosenkranz, G. H. Genzel, C. S. Stîngu, W. Schellenberger, S. Schulz-Stübner, A. C. Rodloff, and K. Eschrich, "Discrimination of human pathogen clostridium species especially of the heterogeneous c. sporogenes and c. botulinum by maldi-tof mass spectrometry," *Current microbiology*, vol. 75, pp. 1506–1515, 2018.

[115] K. Bilecen, G. Yaman, U. Ciftci, and Y. R. Laleli, "Performances and reliability of bruker microflex lt and vitek ms maldi-tof mass spectrometry systems for the identification of clinical microorganisms," *BioMed research international*, vol. 2015, no. 1, p. 516410, 2015.

[116] Bruker Daltonics (16 October 2024, last accessed date). [Online]. Available: https://www.bruker.com/

[117] BioMérieux (16 October 2024, last accessed date). [Online]. Available: https://www.biomerieux.com/

[118] B. Daltonics, "flexanalysis 3.4 user manual," 2011.

[119] R. Ketterlinus, S.-Y. Hsieh, S.-H. Teng, H. Lee, and W. Pusch, "Fishing for biomarkers: analyzing mass spectrometry data with the new clinprotools™ software," *Biotechniques*, vol. 38, no. S6, pp. S37–S40, 2005.

[120] A. S. Lee, H. De Lencastre, J. Garau, J. Kluytmans, S. Malhotra-Kumar, A. Peschel, and S. Harbarth, "Methicillin-resistant staphylococcus aureus," *Nature reviews Disease primers*, vol. 4, no. 1, pp. 1–23, 2018.

[121] N. A. Turner, B. K. Sharma-Kuinkel, S. A. Maskarinec, E. M. Eichenberger, P. P. Shah, M. Carugati, T. L. Holland, and V. G. Fowler Jr, "Methicillin-resistant staphylococcus aureus: an overview of basic and clinical research," *Nature Reviews Microbiology*, vol. 17, no. 4, pp. 203–218, 2019.

[122] G. Daikos and A. Markogiannakis, "Carbapenemase-producing klebsiella pneumoniae:(when) might we still consider treating with carbapenems?" *Clinical Microbiology and Infection*, vol. 17, no. 8, pp. 1135–1141, 2011.

[123] A. E. Stapleton, F. M. Wagenlehner, A. Mulgirigama, and M. Twynholm, "Escherichia coli resistance to fluoroquinolones in community-acquired uncomplicated urinary tract infection in women: a systematic review," *Antimicrobial agents and chemotherapy*, vol. 64, no. 10, pp. e00 862–20, 2020.

[124] B. D. Johnston, P. Thuras, S. B. Porter, M. Anacker, B. VonBank, P. S. Vagnone, M. Witwer, M. Castanheira, and J. R. Johnson, "Global molecular epidemiology of carbapenem-resistant escherichia coli (2002–2017)," *European Journal of Clinical Microbiology & Infectious Diseases*, pp. 1–13, 2021.

[125] A. Elshamy and K. Aboshanab, "A review on bacterial resistance to carbapenems: epidemiology, detection and treatment options," *Future Science OA*, vol. 6, 01 2020.

[126] M. F. Kilkenny and K. M. Robinson, "Data quality:"garbage in–garbage out"," pp. 103–105, 2018.

[127] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *Journal of Computer and System Sciences*, vol. 55, no. 1, pp. 119–139, 1997.

[128] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '16. New York, NY, USA: ACM, 2016, pp. 785–794. [Online]. Available: http://doi.acm.org/10.1145/2939672.2939785

[129] C. Sevilla-Salcedo, A. Guerrero-López, P. M. Olmos, and V. Gómez-Verdejo, "Bayesian sparse factor analysis with kernelized observations," *Neurocomputing*, vol. 490, pp. 66–78, 2022.

[130] Python Software Foundation, *Python Language Reference, version 3*, 2023, available at http://www.python.org.

[131] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[132] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, "TensorFlow: Large-scale machine learning on heterogeneous systems," 2015, software available from tensorflow.org. [Online]. Available: https://www.tensorflow.org/

[133] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2021. [Online]. Available: https://www.R-project.org/

[134] T. M. Inc., "Matlab version: 24.1.0 (r2024a)," Natick, Massachusetts, United States, 2024. [Online]. Available: https://www.mathworks.com

[135] Applied Maths NV, *Bionumerics v7.6.3*, 2024, sint-Martens-Latem, Belgium. [Online]. Available: https://www.applied-maths.com

[136] ASTA, *AMRQuest Software, v.2.1*, 2024, suwon, Korea. [Online]. Available: https://www.astasolutions.com

[137] Clover Bioanalytical Software, *Clover: MALDI-TOF MS Data Analysis Software*, Clover Bioanalytical Software, Granada, Spain, 2023, available at https://www.cloverms.com.

[138] GitHub, Inc., "GitHub," https://github.com, 2023.