Research Directions: Biotechnology Design

www.cambridge.org/btd

Results

Cite this article: Ding Y and Chen S (2025). A bibliographic outlook: machine learning on biofilm. *Research Directions: Biotechnology Design*. **3**, e2, 1–7. https://doi.org/10.1017/btd.2024.28

Received: 9 September 2024
Revised: 25 November 2024
Accepted: 13 December 2024

Keywords:
Biofilm; machine learning; bibliographic analysis; Web of Science; VOSviewer

Corresponding author:
Shan Chen; Email: chen.shan@nie.edu.sg

A bibliographic outlook: machine learning on biofilm

Yuanzhao Ding[1] and Shan Chen[2]

[1]School of Geography and the Environment, University of Oxford, Oxford, UK and [2]Science of Learning in Education Centre, National Institute of Education, Nanyang Technological University, Singapore

Abstract

A biofilm refers to an intricate community of microorganisms firmly attached to surfaces and enveloped within a self-generated extracellular matrix. Machine learning (ML) methodologies have been harnessed across diverse facets of biofilm research, encompassing predictions of biofilm formation, identification of pivotal genes and the formulation of novel therapeutic approaches. This investigation undertook a bibliographic analysis focused on ML applications in biofilm research, aiming to present a comprehensive overview of the field's current status. Our exploration involved searching the Web of Science database for articles incorporating the term "machine learning biofilm," leading to the identification and analysis of 126 pertinent articles. Our findings indicate a substantial upswing in the publication count concerning ML in biofilm over the last decade, underscoring an escalating interest in deploying ML techniques for biofilm investigations. The analysis further disclosed prevalent research themes, predominantly revolving around biofilm formation, prediction and control. Notably, artificial neural networks and support vector machines emerged as the most frequently employed ML techniques in biofilm research. Overall, our study furnishes valuable insights into prevailing trends and future trajectories within the realm of ML applied to biofilm research. It underscores the significance of collaborative efforts between biofilm researchers and ML experts, advocating for interdisciplinary synergy to propel innovation in this domain.

Introduction

Biofilm constitutes a complex and intriguing community of microorganisms characterized by adhesion to surfaces and the secretion of a self-generated extracellular matrix, denoted as a biofilm matrix (Flemming and Wingender 2010). This matrix serves as a protective barrier, endowing the biofilm with resilience against antibiotics, immune system responses and various environmental factors (Dufour et al. 2010). Naturally widespread, biofilms can colonize diverse surfaces (Flemming and Wuertz 2019), including medical implants, water distribution systems and food processing equipment (Galie et al. 2018), thereby presenting challenges such as infections, biofouling and corrosion, making them a prominent subject of scientific investigation (Wang et al. 2020).

In recent years, there has been an increasing interest in employing machine learning (ML) techniques in biofilm research (Artini et al. 2022; Artini et al. 2018; Papa et al. 2020; Patsilinakos et al. 2019). ML, a subset of artificial intelligence, empowers computer systems to learn and enhance performance through experience, without explicit programming (Lavallin and Downs 2021). Through ML algorithms, the copious and intricate datasets derived from biofilm research, encompassing genomics, proteomics and metabolomics data, can be scrutinized to identify pivotal genes, proteins and metabolites associated with biofilm formation and function (Johnson et al. 2016).

The primary aim of this research is to conduct an exhaustive bibliometric analysis of ML's application in biofilm research, offering insights into the current landscape of the field (Li et al. 2023). Bibliometrics entails a quantitative examination of scientific publications, providing valuable insights into the structure, dynamics and trends of a specific research area (Anwar et al. 2022; Talafidaryani et al. 2023). By employing bibliographic analysis, one can discern influential publications, authors and institutions in the ML-biofilm domain, along with prevalent research topics and trends (Zhang et al. 2017).

Such analysis can also unveil how ML has contributed to advancing our comprehension of biofilm formation, growth and function. The amalgamation of ML and biofilm research has enabled the analysis of extensive and intricate datasets, identifying critical factors influencing biofilm aspects such as formation, growth and function, including environmental influences, genetic composition and metabolic processes. Moreover, ML techniques have potential applications in devising novel strategies for biofilm detection, prevention and control, thereby serving as a valuable tool in addressing challenges associated with biofilm-related issues.

This study will entail a meticulous search of pertinent databases for articles published with the keywords "machine learning biofilm." Focusing on this relatively recent timeframe will provide insights into the most current research trends. Subsequently, the use of bibliographic analysis tools, such as VOSviewer, will facilitate the examination of identified articles, generating bibliographic networks to unveil crucial information about the field's structure and dynamics. The analysis will encompass co-authorship analysis, keyword co-occurrence analysis and citation analysis, among other aspects, to pinpoint influential authors, institutions and publications, as well as common research topics and trends.

In summary, this study aims to offer a comprehensive overview of the current status of ML in biofilm research. By identifying key research themes, trends and influential entities in the field, it seeks to enhance understanding of challenges and opportunities linked to the integration of ML and biofilm research. Furthermore, the analysis is poised to guide future research endeavors, contributing to the advancement of knowledge regarding the intricate biological processes involved in biofilm formation, growth and function.

Materials and methods

To comprehensively assess the contemporary research landscape concerning the utilization of ML in the domain of biofilm, an exhaustive search was conducted on the Web of Science. The aim was to ensure the timeliness and relevance of our analysis. In January 2024, the search query "machine learning biofilm" identified a total of 126 articles that met our specified inclusion criteria (AlRyalat et al. 2019; Archambault et al. 2009; Gorraiz and Schloegl 2008). For an in-depth exploration of the predominant themes and trends within this body of literature, VOSviewer (version 1.6.17 and 1.6.20), a robust software tool for constructing and visualizing bibliographic networks, was employed (Ji et al. 2021; Ramírez-Malule et al. 2020; Shah et al. 2020). Various bibliographic analyses, encompassing co-occurrence analysis, country/region analysis and institution analysis, were executed (Zhu et al. 2020).

The co-authorship analysis focused on discerning the most impactful authors in the intersection of ML and biofilm (Wang et al. 2023). This scrutiny revealed several highly productive authors who have significantly shaped the field. In the co-occurrence analysis, the prevalent research topics and themes within the literature were identified (Xue et al. 2021). The analysis underscored key themes, such as biofilm formation, biofilm detection and the application of ML for identifying and predicting bacterial growth patterns. Furthermore, subtopics within these overarching themes, such as the influence of various environmental factors on biofilm formation and the formulation of ML algorithms for precise prediction of bacterial growth, were also pinpointed (Hashemi et al. 2018).

The citation analysis shed light on the most influential publications and highly cited articles in the ML and biofilm research domain (Rickert et al. 2021). Numerous articles garnered substantial citations, indicating their noteworthy impact on the field. Notably, a considerable portion of these highly cited articles centered around the use of ML for predicting biofilm formation and the development of novel algorithms to enhance the comprehension of bacterial growth patterns.

Our bibliographic analysis furnishes valuable insights into the current status of ML applications in the realm of biofilm research (Li et al. 2023). By identifying influential authors, institutions and publications, as well as unveiling prevalent research themes and trends, we enhance our understanding of the challenges and opportunities in this dynamic field of study (Qi et al. 2019; Zhang et al. 2020). Moreover, these findings can guide future research endeavors, contributing to the advancement of our comprehension of the intricate biological processes involved in biofilm formation and growth (Colares et al. 2020; Moura et al. 2016).

Results

The results of our analysis reveal a significant and accelerating interest in the application of ML techniques to biofilm research over the past decade. Notably, the number of publications in this intersection has shown a rapid increase, indicating a growing recognition of the potential of ML in advancing biofilm-related studies.

The primary focus of research in this area revolves around biofilm formation, prediction and control, showcasing the diverse applications of ML techniques in addressing crucial aspects of biofilm-related processes. Among the various ML techniques employed, artificial neural networks (de Ramón-Fernández et al. 2020; Lahiri et al. 2021; Lesnik and Liu 2017) and support vector machines (Li et al. 2022; Modak et al. 2022; Shengxian et al. 2012) emerge as the most frequently used, underscoring their effectiveness in biofilm research.

Authorship patterns demonstrate a diverse and influential collaboration across academic and research institutions, as well as industrial and commercial organizations. Similarly, influential institutions contributing to this field encompass universities, research centers and hospitals. Noteworthy publications with high citation rates are centered on the development of ML algorithms for predicting biofilm formation and identifying key genes associated with biofilm formation.

The central portion of Figure 1 visually encapsulates the most significant words in the field, highlighting "machine learning" and "biofilm" as central themes. Additionally, critical terms related to biofilm formation, including "*Staphylococcus aureus*," "expression," "adaptation," "motility," "growth" and "virulence," provide insights into key bacterial processes and study targets in biofilm research. Words associated with machine learning, such as "classification," "identification," "algorithm" and "system," underscore the focus on predicting biofilm formation – a central goal in ML applications in this domain. In summary, Figure 1 offers a comprehensive overview of the essential words in the field, emphasizing their interconnectedness and providing valuable insights into the prominent themes in biofilm research and ML applications.

As illustrated in Figure 2, there exists a widespread and robust global interest in research endeavors related to the subject matter under examination. This heightened global interest has culminated in extensive collaborations among various countries and regions, fostering a dynamic and interconnected landscape of knowledge exchange. The visualization notably underscores the substantial involvement of the United States and China, both emerging as pivotal players with a pronounced level of engagement and substantial contributions to the expanding body of knowledge in this field.

While the United States and China take center stage, it is crucial to acknowledge the active participation of several other countries in shaping the research landscape. Countries such as Canada, the United Kingdom, Japan, India, Spain, Germany, Switzerland, Italy, Sweden, Denmark and South Korea have emerged as significant

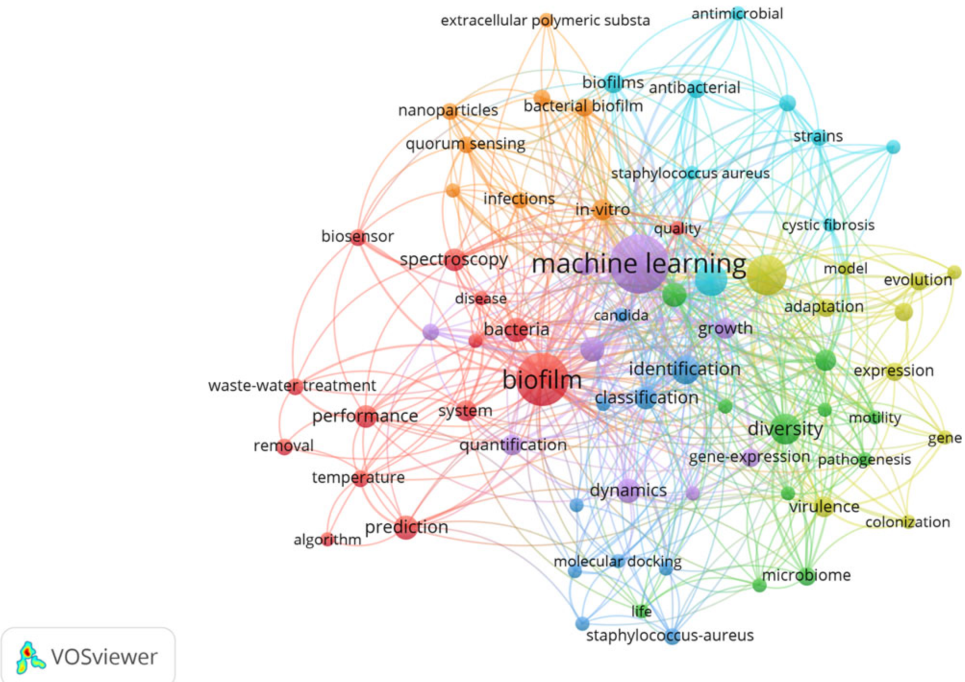

Figure 1. Most important words in this research field and their connection by VOSviewer.

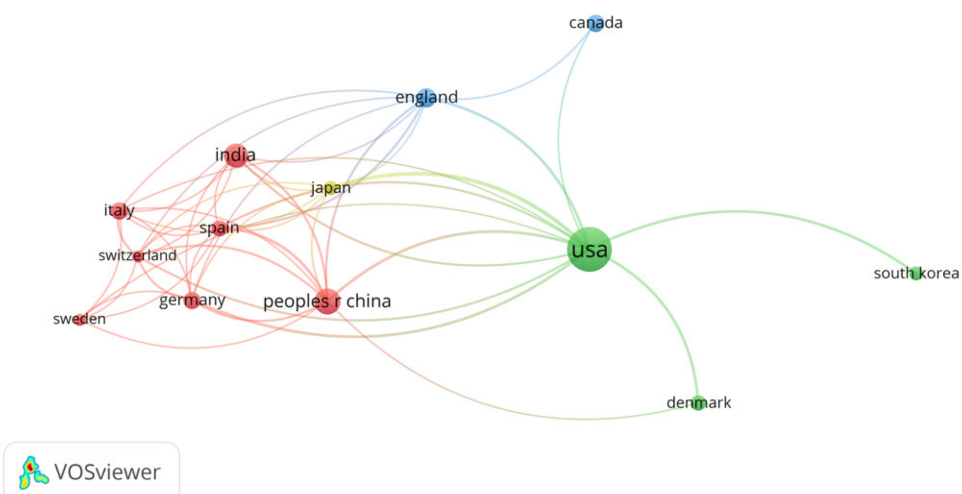

Figure 2. Scientific collaboration network across different countries visualized using VOSviewer. Larger circles represent a higher number of published papers, while connecting lines indicate research collaborations.

contributors to the field, each making noteworthy strides in advancing research frontiers. The multifaceted contributions from this diverse set of countries have enriched the field by introducing varied perspectives and methodological approaches, resulting in a more nuanced and comprehensive understanding of the subject matter.

Collaborative endeavors and cross-cultural exchange of ideas have proven instrumental in propelling advancements within this domain. The contributions of each participating country, with its unique research strengths and approaches, have collectively contributed to a richer and more intricate tapestry of knowledge. This synergy has not only accelerated progress but has also fostered a more inclusive and globally informed research community.

As evidenced by the current landscape, the importance of global collaborations in driving innovation and addressing complex challenges is unmistakable. The ongoing trend of international cooperation in research activities is anticipated to persist and even intensify in the future. This collaborative spirit promises to further amplify the impact of research endeavors, offering a collective and diverse approach to tackling the multifaceted dimensions of the subject under investigation. In essence, the interconnected and collaborative nature of global research efforts is poised to play a pivotal role in shaping the future trajectory of advancements in this field.

As delineated by the insights gleaned from Figure 3, a comprehensive analysis of the institutions at the forefront of research in the intersection of ML and biofilm reveals a diverse and influential array of academic and medical research entities, each playing a significant role in shaping and advancing the discourse within this burgeoning field.

The University of Mississippi stands as a standout institution, positioned as a keystone in the research landscape of ML on biofilm. Renowned for its academic prowess and research excellence, the University of Mississippi has consistently demonstrated a commitment to pushing the boundaries of knowledge within this specific

https://doi.org/10.1017/btd.2024.28 Published online by Cambridge University Press

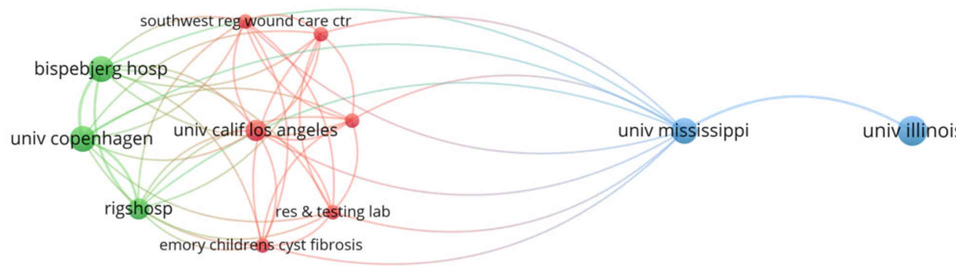

Figure 3. Scientific collaboration network across different organizations visualized using VOSviewer. Larger circles represent a higher number of published papers, while connecting lines indicate research collaborations.

VOSviewer

domain. Its contributions underscore a dedication to fostering innovation and breakthroughs, establishing it as a pivotal player in the ongoing pursuit of advancements.

Another prominent player in this research arena is the University of Illinois, celebrated for its dedication to cutting-edge research initiatives at the confluence of ML and biofilm studies. With a reputation for rigorous inquiry and a proclivity for innovative approaches, the University of Illinois has solidified its standing as a catalyst for advancements, creating an environment conducive to scholarly endeavors within this field.

The University of California at Los Angeles (UCLA) emerges as a beacon of academic and research excellence, contributing significantly to the vibrancy of research activities within the dynamic field of ML on biofilm. Renowned for its research acumen and unwavering commitment to scientific exploration, UCLA plays a crucial role in furthering our understanding of the nuanced intersections between ML and biofilm studies.

In the medical research sphere, the Southwest Regional Wound Care Center, Bispebjerg Hospital, Rigshospitalet and Emory Children's Cystic Fibrosis Research and Testing Lab collectively represent a convergence of institutions, each offering a unique and specialized perspective on the dynamic landscape of ML on biofilm. These medical institutions contribute to the holistic understanding of the subject, providing valuable insights and advancements in the field of medical research, particularly in the context of biofilm studies.

The global reach of research is further highlighted by the inclusion of international institutions such as the University of Copenhagen. These institutions add an international dimension to the collaborative efforts, contributing diverse perspectives and methodologies that enrich the global discourse and contribute to the comprehensive understanding of ML applications in biofilm research.

Collectively, this assembly of diverse and influential institutions showcased in Figure 3 underscores the collaborative and multidisciplinary nature of research endeavors at the intersection of ML and biofilm. The inclusion of institutions from various geographical locations and academic disciplines not only enriches the global research community but also underscores the interconnectedness of efforts aimed at addressing the multifaceted challenges inherent in this specific research area.

The institutions highlighted in Figure 3 collectively constitute a nexus of academic and medical excellence in the dynamic realm of ML on biofilm. Their combined endeavors, marked by scholarly rigor and a collaborative spirit, contribute significantly to the vibrancy and dynamism of research activities in this burgeoning field. This collaborative ecosystem emphasizes the importance of a global network of institutions working synergistically to advance our understanding of the subject matter and collectively contribute to the overarching goals of scientific exploration in ML applications for biofilm research.

Discussion

Revolutionizing biofilm management: the integration of machine learning techniques

When studying bacterial biofilms, previous research has primarily relied on experiments and models. These models are categorized into traditional predictive models and the ML models discussed in this paper. In the quest to develop effective strategies for managing bacterial biofilms, scientists have consistently explored the potential of ML models. Table 1 presents a comparison between ML models and traditional models.

Upon scrutinizing Table 2, which delineates recent and seminal papers in the field, it becomes apparent that ML techniques have been extensively employed in the study of biofilms. Notably, a significant study conducted by a group of researchers utilized an ML model to predict the presence of biofilm inhibitory molecules (Srivastava et al. 2020). This investigation incorporated a combination of descriptor, fingerprint and hybrid models, achieving impressive accuracy rates of 93%, 88% and 90%, respectively. Furthermore, the software resulting from this study, Molib, has evolved into a widely utilized tool for predicting small molecules with biofilm inhibitory properties. The implications of the success of Molib are particularly promising, offering an opportunity for therapeutic intervention against bacteria capable of forming biofilms.

One notable investigation involved the use of *Pseudomonas aeruginosa*, a commonly studied model organism, to scrutinize the chemical components of essential oils (EOs) and their potential impact on biofilm formation (Artini et al. 2022; Artini et al. 2018). In this study, the researchers employed 11 different classification models (F1–F11) to analyze the data and assess the accuracy of the ML predictions. The results demonstrated that the models achieved prediction accuracies ranging from 69% to 98%, underscoring the efficacy of ML in identifying EO chemical components that may impact biofilm formation. Through their analysis, the authors pinpointed specific EO chemical components potentially influencing bacterial biofilm formation in both positive and negative ways. This insight is crucial for scientists working on developing strategies for the effective management and prevention of potentially harmful biofilms.

https://doi.org/10.1017/btd.2024.28 Published online by Cambridge University Press

Table 1. Comparison between traditional prediction models versus machine learning (ML) models

	Traditional prediction model	ML models
Rational	Prediction based on experimental and empirical knowledge	Predictions made using big data connections and ML models
Methods	A model based on experiments and scientists' understanding of bacteria	A model based on big data and ML algorithms
Determinants of accuracy	The accuracy of scientists' logic link (e.g., richer media yields more biofilm)	The quantity and quality of data
Advantages	A well-established method used for many years	A rapid and novel method quickly providing predictive results
Limitations	Flawed logic link leading to inaccurate predictive model outcomes	Black swan events (unseen in the database before) causing inaccurate predictions
References	(Guggenheim et al. 2004; McBain 2009)	(Chen and Ding 2022; Patsilinakos et al. 2019)

Table 2. Summary of recent and important biofilm machine learning (ML) studies

Model organism	Target/biofilm process	ML models	ML accuracy	Main contributions	References
Pseudomonas aeruginosa	Essential oil chemical components	Binary Classification	69%–98%	ML to identify chemical components responsible for bacterial biofilm formation	(Artini et al. 2022; Artini et al. 2018)
Staphylococcus aureus and Staphylococcus epidermidis	Essential oil chemical components	Binary classification	68.7%–90.6%	ML to identify chemical components that modulate biofilm production	(Patsilinakos et al. 2019)
S. aureus	Essential oil chemical components	Binary classification	NA	ML to predict essential oils modulate biofilm production and inhibit biofilm	(Papa et al. 2020)
S. aureus	Acyl-CoA thioesterase	Classification	59.46–94.59%	Identification of 36 candidate genes including an acyl-CoA thioesterase enzyme and ten hypothetical proteins	(Subramanian and Natarajan 2021)
S. aureus, P. aeruginosa, Acinetobacter baumannii, Stenotrophomonas maltophilia, Escherichia coli	Biofilm infection	Random forest	95.0%–100%	Using lanthanide nanoparticles detects pathogenic biofilms based on random forest	(Wang et al. 2022)

Another article discussed the challenges in treating biofilm-associated infections caused by *Staphylococcus aureus* and *Staphylococcus epidermidis* (Patsilinakos et al. 2019). The study investigated the potential of EOs as a treatment option and examined the ability of 89 EOs to influence biofilm production in various bacterial strains. ML algorithms analyzed the chemical compositions of the EOs to evaluate their anti-biofilm potencies and pinpoint the components responsible for biofilm production, inhibition or stimulation.

In another study, EOs were investigated as natural alternatives to chemotherapeutic drugs for inhibiting biofilm in chronic *S. aureus* infections (Papa et al. 2020). A total of 61 EOs were evaluated for biofilm modulation and antibacterial activity. Their chemical composition was analyzed using GC/MS, and ML algorithms were employed to correlate potency with active components. Certain EOs inhibited biofilm growth at a 1.00% concentration and were further characterized for their effects on biofilm organization through scanning electron microscope studies.

Another paper presented a novel computational methodology that combines meta-analysis and ML to identify important genes and pathways in biofilm-forming bacteria (Subramanian and Natarajan 2021). This approach analyzed gene expression profiles in multiple *S. aureus* strains and identified 36 potential genes, including 11 newly reported ones. These genes are considered essential for biofilm development and represent a signature target list for designing anti-biofilm therapeutics. The study underscores the value of combining meta-analysis and ML techniques to enhance understanding of biofilm mechanisms and advance effective therapeutic strategies.

Another study developed a machine-learning-aided cocktail assay (Wang et al. 2022). This study utilized lanthanide nanoparticles with varied properties, integrated into cocktail kits. The physicochemical diversity of biofilms was translated into luminescence intensity, enabling identification of unknown biofilms with an overall accuracy rate surpassing 80% via the random forest algorithm. Antibiotic-loaded cocktail nano-probes effectively eradicated biofilms, demonstrating the technique's promise as a reliable diagnostic tool for biofilm infections. Moreover, this approach offers a framework for developing assays to detect biochemical compounds beyond biofilms.

Interdisciplinary synergy: biofilm researchers and ML experts shaping the future

Our bibliometric analysis offers a comprehensive panorama of the current landscape of ML in biofilm research. The examination underscores the escalating interest in applying ML techniques to biofilm research and underscores the pivotal role of interdisciplinary collaboration between biofilm researchers and ML experts in propelling innovation within this domain (Hashemi et al. 2018).

The predominant research themes in this realm revolved around biofilm formation, prediction and control, signifying the pressing demand for novel strategies to address biofilm-related challenges (Alaoui Mdarhri et al. 2022). Notably, artificial neural networks (de Ramón-Fernández et al. 2020; Lahiri et al. 2021; Lesnik and Liu 2017) and support vector machines (Li et al. 2022; Modak et al. 2022; Shengxian et al. 2012) emerged as the most frequently employed ML techniques in biofilm research, suggesting their aptness for deciphering intricate biofilm-related data.

An exploration of authorship and institutional affiliations revealed a diverse array of influential authors and institutions spanning various fields, underscoring the interdisciplinary essence inherent in biofilm research. This diversity assumes paramount importance in fostering innovation within the field.

Future recommendation of using ML in bacterial and biofilm studies

The utilization of big data and ML techniques has witnessed a growing prevalence across diverse fields, encompassing species distribution (Chen and Ding 2022; Gobeyn et al. 2019), education (Chen and Ding 2023) and cancer prediction (Cammarota et al. 2020). ML offers a platform through which policymakers can adjust policies to better serve the populace. However, despite the abundance of studies on bacteria and biofilm, the application of ML in this domain remains limited, as elucidated in the preceding paragraphs.

Bacteria and biofilm stand as focal points of extensive investigation within various environmental and industrial contexts, such as heavy metal pollutant removal (Ding et al. 2014) and microbial fuel cells (Zhao et al. 2015). Scientists have developed methods to genetically modify bacterial genes in efforts to either enhance or weaken biological processes, resulting in the formation of stronger or weaker biofilms. Despite these advancements, the underlying mechanisms governing these biological processes remain shrouded in uncertainty (Tribedi et al. 2015).

Given the diverse applications and significance of bacteria and biofilm research, it becomes imperative to explore the potential advantages of incorporating ML techniques into this realm (Sadeghi et al. 2022). By harnessing the voluminous data generated through research endeavors, ML stands poised to unravel the intricate interactions and mechanisms inherent in bacterial processes, offering novel insights and predictions (Long et al. 2022). Furthermore, the development of new ML algorithms specifically tailored to the unique challenges posed by bacterial data holds the promise of delivering more accurate and efficient results (Cordier et al. 2019; Long et al. 2021).

In essence, the integration of ML into bacteria and biofilm research carries the potential to propel our comprehension of these vital biological processes, potentially leading to groundbreaking discoveries and applications across various fields.

Conclusions

Our comprehensive investigation into the integration of ML methodologies in biofilm research reveals a significant surge in interest over the last decade. The bibliographic analysis, encompassing 126 pertinent articles from the Web of Science database, sheds light on prevailing trends, particularly focusing on biofilm formation, prediction and control. Artificial neural networks and support vector machines emerged as the predominant ML techniques. The study underscores the vital role of collaborative efforts between biofilm researchers and ML experts, emphasizing interdisciplinary synergy. The findings provide valuable insights into the current landscape and future trajectories of ML in biofilm research, guiding further exploration and innovation in this dynamic and crucial field.

Data availability statement. Data availability does not apply to this article.

Author contribution. YD conceived and designed the study. YD conducted data gathering. SC performed software. SC performed visualization. YD wrote the article.

Financial support. This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

Competing interests. The authors have no conflicts of interest to declare for this publication.

Ethics statement. Ethical approval and consent are not relevant to this article type.

Connections references

Carlström V, Rigobello A and Ayres P (2023). Modelling of morphogenesis to support the design of fungal-based engineered living materials. *Research Directions: Biotechnology Design* **2**, e11, 1–16. https://doi.org/10.1017/btd.2024.10

References

Alaoui Mdarhri H, Benmessaoud R, Yacoubi H, Seffar L, Guennouni Assimi H, Hamam M, Boussettine R, Filali-Ansari N, Lahlou FA and Diawara I (2022) Alternatives therapeutic approaches to conventional antibiotics: advantages, limitations and potential application in medicine. *Antibiotics* **11**, 12, 1826.

AlRyalat SAS, Malkawi LW and Momani SM (2019) Comparing bibliometric analysis using PubMed, Scopus, and Web of Science databases. *JoVE (Journal of Visualized Experiments)* **152**, e58494.

Anwar MA, Zhang Q, Asmi F, Hussain N, Plantinga A, Zafar MW and Sinha A (2022) Global perspectives on environmental Kuznets curve: A bibliometric review. *Gondwana Research* **103**, 135–145.

Archambault É, Campbell D, Gingras Y and Larivière V (2009) Comparing bibliometric statistics obtained from the Web of Science and Scopus. *Journal of the American Society for Information Science and Technology* **60**, 7, 1320–1326.

Artini M, Papa R, Sapienza F, Božović M, Vrenna G, Tuccio Guarna Assanti V, Sabatino M, Garzoli S, Fiscarelli EV and Ragno R (2022) Essential Oils Biofilm Modulation Activity and Machine Learning Analysis on *Pseudomonas aeruginosa* Isolates from Cystic Fibrosis Patients. *Microorganisms* **10**, 5, 887.

Artini M, Patsilinakos A, Papa R, Božović M, Sabatino M, Garzoli S, Vrenna G, Tilotta M, Pepi F and Ragno R (2018) Antimicrobial and antibiofilm activity and machine learning classification analysis of essential oils from different Mediterranean plants against *Pseudomonas aeruginosa*. *Molecules* **23**, 2, 482.

Cammarota G, Ianiro G, Ahern A, Carbone C, Temko A, Claesson MJ, Gasbarrini A and Tortora G (2020) Gut microbiome, big data and machine learning to promote precision medicine for cancer. *Nature Reviews Gastroenterology & Hepatology* **17**, 10, 635–648.

Chen S and Ding Y (2022) Machine learning and its applications in studying the geographical distribution of ants. *Diversity* **14**, 9, 706.

Chen S and Ding Y (2023) A Machine Learning Approach to Predicting Academic Performance in Pennsylvania's Schools. *Social Sciences* **12**, 3, 118.

Colares GS, Dell'Osbel N, Wiesel PG, Oliveira GA, Lemos PHZ, da Silva FP, Lutterbeck CA, Kist LT and Machado ÊL (2020) Floating treatment wetlands: A review and bibliometric analysis. *Science of the Total Environment* **714**, 136776.

Cordier T, Lanzén A, Apothéloz-Perret-Gentil L, Stoeck T and Pawlowski J (2019) Embracing environmental genomics and machine learning for routine biomonitoring. *Trends in Microbiology* **27**, 5, 387–397.

de Ramón-Fernández A, Salar-García MJ, Fernández DR, Greenman J and Ieropoulos IA (2020) Evaluation of artificial neural network algorithms for predicting the effect of the urine flow rate on the power performance of microbial fuel cells. *Energy* **213**, 118806.

Ding Y, Peng N, Du Y, Ji L and Cao B (2014) Disruption of putrescine biosynthesis in *Shewanella oneidensis* enhances biofilm cohesiveness and performance in Cr (VI) immobilization. *Applied and Environmental Microbiology* **80**, 4, 1498–1506.

Dufour D, Leung V and Lévesque CM (2010) Bacterial biofilm: structure, function, and antimicrobial resistance. *Endodontic Topics* **22**, 1, 2–16.

Flemming H-C and Wingender J (2010) The biofilm matrix. *Nature Reviews Microbiology* **8**, 9, 623–633.

Flemming H-C and Wuertz S (2019) Bacteria and archaea on Earth and their abundance in biofilms. *Nature Reviews Microbiology* **17**, 4, 247–260.

Galie S, García-Gutiérrez C, Miguélez EM, Villar CJ and Lombó F (2018) Biofilms in the food industry: health aspects and control methods. *Frontiers in Microbiology* **9**, 898.

Gobeyn S, Mouton AM, Cord AF, Kaim A, Volk M and Goethals PLM (2019) Evolutionary algorithms for species distribution modelling: A review in the context of machine learning. *Ecological Modelling* **392**, 179–195.

Gorraiz J and Schloegl C (2008) A bibliometric analysis of pharmacology and pharmacy journals: Scopus versus Web of Science. *Journal of Information Science* **34**, 5, 715–725.

Guggenheim B, Guggenheim M, Gmür R, Giertsen E and Thurnheer T (2004) Application of the Zürich biofilm model to problems of cariology. *Caries Research* **38**, 3, 212–222.

Hashemi SJ, Bak N, Khan F, Hawboldt K, Lefsrud L and Wolodko J (2018) Bibliometric analysis of microbiologically influenced corrosion (MIC) of oil and gas engineering systems. *Corrosion* **74**, 4, 468–486.

Ji B, Zhao Y, Vymazal J, Mander Ü, Lust R and Tang C (2021) Mapping the field of constructed wetland-microbial fuel cell: A review and bibliometric analysis. *Chemosphere* **262**, 128366.

Johnson CH, Ivanisevic J and Siuzdak G (2016) Metabolomics: beyond biomarkers and towards mechanisms. *Nature Reviews Molecular Cell Biology* **17**, 7, 451–459.

Lahiri D, Nag M, Sarkar T, Dutta B and Ray RR (2021) Antibiofilm activity of α-amylase from Bacillus subtilis and prediction of the optimized conditions for biofilm removal by response surface methodology (RSM) and artificial neural network (ANN). *Applied Biochemistry and Biotechnology* **193**, 1853–1872.

Lavallin A and Downs JA (2021) Machine learning in geography–Past, present, and future. *Geography Compass* **15**, 5, e12563.

Lesnik KL and Liu H (2017) Predicting microbial fuel cell biofilm communities and bioreactor performance using artificial neural networks. *Environmental Science & Technology* **51**, 18, 10881–10892.

Li P, Tong X, Wang T, Wang X, Zhang W, Qian L, Liao J, Diao W, Zhou J and Wu W (2023) Biofilms in wound healing: A bibliometric and visualised study. *International Wound Journal* **20**, 2, 313–327.

Li X, Chen S, Zhang J, Yu L, Chen W and Zhang Y (2022) Optimization of Ultrasonic-Assisted Extraction of Active Components and Antioxidant Activity from Polygala tenuifolia: A Comparative Study of the Response Surface Methodology and Least Squares Support Vector Machine. *Molecules* **27**, 10, 3069.

Long F, Fan J, Xu W and Liu H (2022) Predicting the performance of medium-chain carboxylic acid (MCCA) production using machine learning algorithms and microbial community data. *Journal of Cleaner Production* **377**, 134223.

Long F, Wang L, Cai W, Lesnik K and Liu H (2021) Predicting the performance of anaerobic digestion using machine learning algorithms and genomic data. *Water Research* **199**, 117182.

McBain AJ (2009) In vitro biofilm models: an overview. *Advances in Applied Microbiology* **69**, 99–132.

Modak S, Lahorkar A and Valadi J (2022) Recent Advances in Applications of Support Vector Machines in Fungal Biology. *Laboratory Protocols in Fungal Biology: Current Methods in Fungal Biology*, 117–136.

Moura LKB, Tapety FI, Mobim M, Lago EC, de LobÃ ES, Leal CMdCL, Santos TC and Monte TL (2016) Bacterial association and oral biofilm formation: A bibliometric analysis. African Journal of Microbiology Research 10, 39, 1654–1661.

Papa R, Garzoli S, Vrenna G, Sabatino M, Sapienza F, Relucenti M, Donfrancesco O, Fiscarelli EV, Artini M and Selan L (2020) Essential oils biofilm modulation activity, chemical and machine learning analysis—Application on *Staphylococcus aureus* isolates from cystic fibrosis patients. *International Journal of Molecular Sciences* **21**, 23, 9258.

Patsilinakos A, Artini M, Papa R, Sabatino M, Božović M, Garzoli S, Vrenna G, Buzzi R, Manfredini S and Selan L (2019) Machine learning analyses on data including essential oil chemical composition and in vitro experimental antibiofilm activities against Staphylococcus species. *Molecules* **24**, 5, 890.

Qi Y, Chen X, Hu Z, Song C and Cui Y (2019) Bibliometric analysis of algal-bacterial symbiosis in wastewater treatment. *International Journal of Environmental Research and Public Health* **16**, 6, 1077.

Ramírez-Malule H, Quinones-Murillo DH and Manotas-Duque D (2020) Emerging contaminants as global environmental hazards. A bibliometric analysis. *Emerging Contaminants* **6**, 179–193.

Rickert CA, Hayta EN, Selle DM, Kouroudis I, Harth M, Gagliardi A and Lieleg O (2021) Machine learning approach to analyze the surface properties of biological materials. *ACS Biomaterials Science & Engineering* **7**, 9, 4614–4625.

Sadeghi M, Panahi B, Mazlumi A, Hejazi MA, Komi DEA and Nami Y (2022) Screening of potential probiotic lactic acid bacteria with antimicrobial properties and selection of superior bacteria for application as biocontrol using machine learning models. *LWT* **162**, 113471.

Shah SHH, Lei S, Ali M, Doronin D and Hussain ST (2020) Prosumption: bibliometric analysis using HistCite and VOSviewer. *Kybernetes* **49**, 3, 1020–1045.

Shengxian C, Yanhui Z, Jing Z and Dayu Y (2012) Experimental study on dynamic simulation for biofouling resistance prediction by least squares support vector machine. *Energy Procedia* **17**, 74–78.

Srivastava GN, Malwe AS, Sharma AK, Shastri V, Hibare K and Sharma VK (2020) Molib: A machine learning based classification tool for the prediction of biofilm inhibitory molecules. *Genomics* **112**, 4, 2823–2832.

Subramanian D and Natarajan J (2021) Integrated meta-analysis and machine learning approach identifies acyl-CoA thioesterase with other novel genes responsible for biofilm development in *Staphylococcus aureus*. *Infection, Genetics and Evolution* **88**, 104702.

Talafidaryani M, Jalali SMJ and Moro S (2023) Tracing the evolution of digitalisation research in business and management fields: Bibliometric analysis, topic modelling and deep learning trend forecasting. *Journal of Information Science*, 01655515221148365.

Tribedi P, Gupta AD and Sil AK (2015) Adaptation of Pseudomonas sp. AKS2 in biofilm on low-density polyethylene surface: an effective strategy for efficient survival and polymer degradation. *Bioresources and Bioprocessing* **2**, 1–10.

Wang H, Christiansen DE, Mehraeen S and Cheng G (2020) Winning the fight against biofilms: the first six-month study showing no biofilm formation on zwitterionic polyurethanes. *Chemical Science* **11**, 18, 4709–4721.

Wang J, Jiang Z, Wei Y, Wang W, Wang F, Yang Y, Song H and Yuan Q (2022) Multiplexed identification of bacterial biofilm infections based on machine-learning-aided lanthanide encoding. *ACS Nano* **16**, 2, 3300–3310.

Wang T, Zhang R, Chen Z, Cao P, Zhou Q and Wu Q (2023) A global bibliometric and visualized analysis of bacterial biofilm eradication from 2012 to 2022. *Frontiers in Microbiology* **14**, 1287964.

Xue J, Reniers G, Li J, Yang M, Wu C and van Gelder P (2021) A bibliometric and visualized overview for the evolution of process safety and environmental protection. *International Journal of Environmental Research and Public Health* **18**, 11, 5985.

Zhang S, Mao G, Crittenden J, Liu X and Du H (2017) Groundwater remediation from the past to the future: A bibliometric analysis. *Water Research* **119**, 114–125.

Zhang T, Yin X, Yang X, Man J, He Q, Wu Q and Lu M (2020) Research trends on the relationship between microbiota and gastric cancer: a bibliometric analysis from 2000 to 2019. *Journal of Cancer* **11**, 16, 4823.

Zhao C, Wu J, Ding Y, Wang VB, Zhang Y, Kjelleberg S, Loo JSC, Cao B and Zhang Q (2015) Hybrid conducting biofilm with built-in bacteria for high-performance microbial fuel cells. *ChemElectroChem* **2**, 5, 654–658.

Zhu Y, Li JJ, Reng J, Wang S, Zhang R and Wang B (2020) Global trends of Pseudomonas aeruginosa biofilm research in the past two decades: A bibliometric study. *MicrobiologyOpen* **9**, 6, 1102–1112.