International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2022): 7.942

Discovering the Use of Machine Learning and Deep Learning for AntiMicrobial Resistance Detection

Srivaramangai R

Department of Information Technology, University of Mumbai, India
Email: *rsrimangai[at]gmail.com*

Abstract: ***A major threat to modern medicine is antimicrobial resistance (AMR), a global health crisis. To halt the emergence and further spread of AMR, effective prevention strategies are urgently required. Clinical microbiology currently primarily uses two methods to diagnose AMR. They are Whole - genome sequencing for antimicrobial susceptibility testing (WGS - AST) and traditional culture - based antimicrobial susceptibility testing (AST). Although the development of molecular tests has significantly improved the speed of diagnostic testing as well as the prompt identification of pathogens and patterns of antibiotic resistance, their high costs and limited availability prevent their widespread use. Machine learning (ML) is increasingly being used to predict pathogen resistance to various antibiotics based on gene content and genome composition given the availability of data sets containing hundreds or thousands of pathogen genomes. Predictive models of antimicrobial resistance that are driven by machine learning (ML) and deep learning (DL) may be able to bridge the gap between the results of molecular and genotypic susceptibility analysis and the collection of specimens, making it easier to choose empirical antibiotics in a faster way. The research project aims to promote the use of machine learning (ML) in front - line settings while also highlighting the additional refinements required to use these methods safely and with confidence. The use of machine learning is not new in this field. Thus, in this paper, the review of such research works is done with comparative analysis and inferences.***

Keywords: AntiMicronial Resistance, Antibiotic Resistance, Machine Learning, Deep Learning, AMPTrans - lstm, XGBoost

1. Introduction

One of modern medicine's most challenging phenomena is still antimicrobial resistance (AMR). AI (ML) is a subfield of man - made rational thinking and logical reasoning that spotlights on the improvement of calculations that figure out how to precisely foresee result factors utilizing enormous arrangements of identifying factors that are regularly not hand chosen and are insignificantly organized. Generally, the predictive performance of models is evaluated by applying them to a test data set and parameterizing them using a training data set. In the past five years, there has been a growing interest in applying machine learning algorithms to the issue of AMR as a result of the exponential growth of clinical and experimental data, significant investments in computational capacity, improvements in algorithm performance, and the growing need for novel strategies to lessen the burden of disease. Further to add in this, treatment optimisations for sexually transmitted infections (STIs) are very much required due to the emergence of antimicrobial resistance in Neisseria gonorrhoeae, Mycoplasma genitalium, and Trichomonas vaginalis. In this paper, a review of existing machine learning and deep learning methods and related works on AMR predictions and suggestions are studied and also compared for its performance. Most of the work has been published in PubMed and the number of papers found in the area of ML and DL is less. So there is a lot of scope for the IT researchers to explore this area to find better performing techniques. This paper is organized as I. Introduction, II. AMR, III. Review of ML/DL Techniques, IV. Conclusion followed by the references.

2. AntiMicrobial Resistance (AMR)

According to WHO, AMR is one of the top 10 global public health threats worldwide which is prominently affected in human beings. When microorganisms like bacteria and fungi acquire the capacity to overcome drugs intended to kill them, this phenomenon is known as antimicrobial resistance. AMR occurs naturally over time, typically as a result of genetic modifications. People, animals, food, plants, and the environment (in water, soil, and air) all contain antimicrobial - resistant organisms. They can spread through food of an animal's origin or between people and animals, as well as between people. The development of drug - resistant pathogens is primarily fuelled by the misuse and overuse of antibiotics. Figure 1 illustrate the reasons for the occurrence and spread of AMR especially AR. So it is obvious that immunity level, use of contaminated food, water, consumption of prolonged medicines, improper diagnosis and prescription etc are the major factors that causes Antibiotic resistance [1]. According to Galhano et. al [2], since it can collect all information about a microbiome that is present in the sample being studied, metagenomics is regarded as a method that is somewhat more comprehensive than other molecular methods in detection process. In terms of AMR determination, however, there is no better or worse method, and the best combination of each method's advantages depends on the question that needs to be answered.

Machine learning algorithms have been in use in various research works on diagnosis or data analysis of genes for AMR. Madrigal et. al [3] have analyzed the data with a deep learning model, by expanding the list of AMR genes and surpass conventional cut - offs based solely on high DNA sequence similarity. The study by Shankarnarayan et. al [4] investigates the role that machine learning plays in the design and discovery of antibacterial and antifungal drugs. According to them, overall, machine learning promises to facilitate the creation of antibacterial and antifungal medications and advance research on antimicrobial resistance. Multimodal antibiotic stewardship programs are essential for limiting antibiotic misuse and preventing further aggravation of the AMR problem as antibiotic efficacy and treatment options decrease. Predicting early antibiotic resistance [5] has been made possible with the help of both supervised and unsupervised machine learning tools, assisting physicians in selecting the most appropriate course of treatment. There are works done where they have combined antimicrobial resistance prediction with existing research on machine learning and artificial intelligence (AI) in general. Despite increased antibiotic resistance, there is mounting evidence that machine learning can assist physicians in selecting the appropriate anti - infective treatment based on patient demographics and clinical histories. The analysis of bacterial genomes, the prediction of medication susceptibility, the identification of epidemic patterns for the purpose of surveillance, and the development of new antibacterial treatments or vaccines are all applications for machine learning (ML) techniques. Since one of the main causes of AMR is the inappropriate use of antibiotics, optimizing the use of existing medications has been a top priority in addition to the development of new antibiotics. So, the next section is focused on reviewing intensely on these ML/DL based research works and to find out how much of success have been achieved and further how much scope is there for future work.

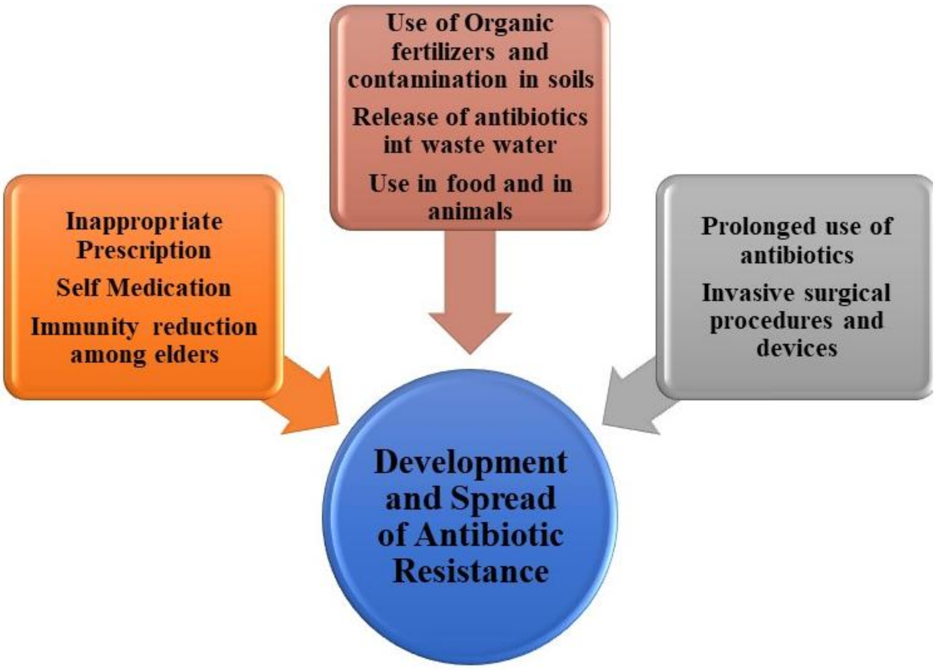

Figure 1: AMR - AR development

3. Review of ML/DL Techniques

Kim et. al [6] in their work have given the variety of quantitative and qualitative laboratory measures of AMR, predicting what to expect is not an easy task. They further have said that ML models treat genes as independent predictors without taking into account their structural and functional connections; When new mutational variants of known AMR genes emerge, they may also not be accurate. So they conclude that ML models need to be transparent and easy to understand in order for end users in public health settings to trust the technology. They further suggest that in order to close the diagnostic implementation gap, the subsequent set of AMR - ML studies can concentrate on improving these limitations. Anahtar et. al [7] gives emphasis on three work areas, and they examine the current state of research at the joining of ML and AMR. The first is using genomic data to predict AMR. The second is the application of machine learning to gain insight into the cellular functions that are disrupted by antibiotics, which serves as the foundation for comprehending the mechanisms of action and the creation of novel antibiotics. The third focuses on using data from the electronic health record to use machine learning for antimicrobial stewardship. According to them, even though ML is still in its infancy when it comes to understanding, diagnosing, treating, and preventing AMR, the growing amount of data and interest in it guarantees that it will become a crucial component of any subsequent translational research programs. Madrigal et. al [3], in their study computational predictions and validation analyses found that the machine learning can be used to find hidden AMR determinants in metagenomics datasets, expanding the knowledge of the ISS environmental microbiomes and their potential to cause disease in humans. Basically, these researchers had worked on a deep learning model which was used to analyze the data, enabling them to expand the list of AMR genes and go beyond conventional cutoffs based solely on high DNA sequence similarity. AMR predominated in the final flight for Kalamiella Piersonii, a human urinary tract infection - causing bacteria, as shown by our PMA - treated samples. Macesic et. al [5], have described and found that machine learning can be used to study AMR, but only in limited ways. There are concerns about model interpretability and data quality pose barriers to the adoption of machine learning in clinical settings. Predicting antimicrobial susceptibility phenotypes and other laboratory - based applications of machine learning to AMR are likely to be the future research areas. They have highlighted that use of machine learning is a significant contribution to the research that could be used to predict resistance in other bacteria and to other antibiotics and can be successfully applied even in complex forms of antibiotic resistance. Sena et. al [8] describes their goal which is to bring out the pressing issues in global STI management such as: 1) enhanced monitoring of antimicrobial resistance and clinical outcomes; 2) systematic pharmacokinetic and pharmacodynamic analyses to guarantee bacterial eradication and suppression of resistance at all infection sites; 3) the creation of novel, cost - effective antibiotics; and (4) advancements in new molecular and point - of - care tests for identifying factors that contribute to antimicrobial resistance. Antimicrobial opposition among STIs is a worldwide general wellbeing emergency. In addition to other public health interventions, ongoing efforts to develop novel antimicrobials will be necessary to lessen the global burden of STIs. A crucial research priority is the development of STI vaccines to prevent transmission, in addition to prevention through safer sexual practices. Macesic et. al [9] in their study again on the same field, have highlighted that use of machine learning is a significant contribution to the research that could be used to predict resistance in other bacteria and to other antibiotics and can be successfully applied even in complex forms of antibiotic resistance. Hicks et. al [10] have shown that WGS data can be utilized by ML to aid in the creation of rapid molecular diagnostics. Further they have found that, in addition to ML methods and parameters, a variety of other factors can have an impact on model performance, reliability, and interpretability, although a larger sampling of methods and parameters will be required to achieve model performance optimization. Backbone capping or lysine - to - arginine substitutions in P1 variants make a bacterial strain that has evolved resistance to this peptide susceptible. This is based on the finding of Huynh et. al [11], where they further add that a cocktail of AMPs that are closely related may be useful for overcoming evolved resistance. A XGBoost - based machine learning model has been used by Nguyen et. al [12] that accurately predicts minimum inhibitory concentrations (MICs) for 20 antibiotics was created using whole genome sequence data from 1668 clinical Klebsiella pneumoniae isolates. The model's overall accuracy is 92 percent within ±1 twofold dilution factor. They found that deeper trees, with a depths of 3–4, were optimal for the XGBoost model. Camp, Haslam and Porollo [13] says that the whole genome sequencing enables the prediction of antimicrobial resistance in Gram - negative bacteria. They have provided an in silico antibiogram for eight drugs using a tool. A reliability index is included with predictions, which may make it easier to make decisions. Moradigaravand et. al [14] used existing and new whole genome sequences of 1936 E. coli strains to predict resistance to 11 compounds from four antibiotic classes using machine learning models. As predictors, they focused at population structure, isolation year, gene content, and polymorphism information using a variety of approaches. On held - out data, gradient boosted decision trees consistently performed better than alternative models, with an average accuracy of 0.91. Her and Wu [15], concluded that based on a genetic algorithm (GA), subsets of AMR genes were chosen for various antibiotics with better prediction results than the established gene sets. This suggests that the gene sets chosen by the GA may need more research to learn more about how E. coli fights antibiotics. Mao et. al [16] presented AMPTrans - lstm, a method for the rational design of AMPs that is based on a deep generative network. Module identification, transfer learning, and pre - training are all components of the AMPTrans - lstm pipeline. The (long short - term memory) LSTM sampler and Transformer converter submodels of the AMPTrans - lstm model can be connected in series to take full advantage of the stability of the LSTM and the novelty of the Transformer model. These components might result in candidates for AMPs, which could then be tailored for particular applications. The researchers have demonstrated that AMPTrans - lstm is capable of producing reasonable and novel AMPs sequences by analyzing the generated sequence and trained AMPs. An effective AMPs design tool, AMPTrans - lstm is capable of producing functional peptides for antimicrobial resistance that are both novel and diverse.

4. Conclusion

The ML and DL based applications are briefly and comprehensively examined in this study. Current techniques can process datasets with a similar conveyance, so their generalizability is inadequate. Algorithms like SVM, Naïve Baise, XGBoost, decision tree, AMpTrans etc have been used including the evolutionary techniques like genetic algorithms. But the redundancy of using these techniques are less and so the performance or heavily used technique is unpredictable at this stage. The accuracy level of these techniques are still not promising. The use of unsupervised learning to automatically annotate unlabeled data is one possible direction for future research. Moreover, detection techniques need to be further enhanced with new algorithms or modifying the existing ones to level up the performance.

References

[1] Prestinaci F, Pezzotti P, Pantosti A. Antimicrobial resistance: a global multifaceted phenomenon. Pathog Glob Health.2015; 109 (7): 309 - 18. doi: 10.1179/2047773215Y.0000000030. Epub 2015 Sep 7. PMID: 26343252; PMCID: PMC4768623.

[2] S P Galhano B, G Ferrari R, Panzenhagen P, S de Jesus AC, A Conte - Junior C. Antimicrobial Resistance Gene Detection Methods for Bacteria in Animal - Based Foods: A Brief Review of Highlights and Advantages. Microorganisms.2021 Apr 26; 9 (5): 923. doi: 10.3390/microorganisms9050923. PMID: 33925810; PMCID: PMC8146338.

[3] Madrigal, P., Singh, N. K., Wood, J. M. et al. Machine learning algorithm to characterize antimicrobial resistance associated with the International Space Station surface microbiome. Microbiome 10, 134 (2022). https: //doi. org/10.1186/s40168 - 022 - 01332 - w.

[4] Shankarnarayan, Shamanth, Joshua D. Guthrie, and Daniel A. Charlebois.2022. “Machine Learning for Antimicrobial Resistance Research and Drug Development. ” The Global Antimicrobial Resistance Epidemic - Innovative Approaches and Cutting - Edge Solutions, September. IntechOpen. doi: 10.5772/intechopen.104841.

[5] Macesic, N.; Polubriaginof, F.; Tatonetti, N. P. Machine learning: Novel bioinformatics approaches for combating antimicrobial resistance. Curr. Opin. Infect. Dis.2017, 30, 511–517.

[6] Kim JI, Maguire F, Tsang KK, Gouliouris T, Peacock SJ, McAllister TA, McArthur AG, Beiko RG. Machine Learning for Antimicrobial Resistance Prediction: Current Practice, Limitations, and Clinical Perspective. Clin Microbiol Rev.2022 Sep 21; 35 (3): e0017921. doi: 10.1128/cmr.00179 - 21. Epub 2022 May 25. PMID: 35612324; PMCID: PMC9491192.

[7] Anahtar MN, Yang JH, Kanjilal S. Applications of Machine Learning to the Problem of Antimicrobial Resistance: an Emerging Model for Translational Research. J Clin Microbiol.2021 Jun 18; 59 (7): e0126020. doi: 10.1128/JCM.01260 - 20. Epub 2021 Jun 18. PMID: 33536291; PMCID: PMC8218744.

[8] Sena AC, Bachmann L, Johnston C, Wi T, Workowski K, Hook EW 3rd, Hocking JS, Drusano G, Unemo M. Optimising treatments for sexually transmitted infections: surveillance, pharmacokinetics and pharmacodynamics, therapeutic strategies, and molecular resistance prediction. Lancet Infect Dis.2020 Aug; 20 (8): e181 - e191. doi: 10.1016/S1473 - 3099 (20) 30171 - 7. Epub 2020 Jun 19. PMID: 32569625; PMCID: PMC8041119.

[9] Macesic N, Walk OIBD, Pe’er I, Tatonetti NP, Peleg AY, Uhlemann AC.2020. Predicting phenotypic polymyxin resistance in *Klebsiella pneumoniae* through machine learning analysis of genomic data. *mSystems* 5: e00656 - 19.

[10] Hicks AL, Wheeler N, Sanchez - Buso L, Rakeman JL, Harris SR, Grad YH.2019. Evaluation of parameters affecting performance and reliability of machine learning - based antibiotic susceptibility testing from whole - genome sequencing data. *PLoS Comput Biol* 15: e1007349.

[11] Huynh L, Velásquez J, Rabara R, Basu S, Nguyen HB, Gupta G. Rational design of antimicrobial peptides targeting Gram - negative bacteria. Comput Biol Chem.2021 Jun; 92: 107475. doi: 10.1016/j. compbiolchem.2021.107475. Epub 2021 Mar 17. PMID: 33813188. Huynh L, Velásquez J, Rabara R, Basu S, Nguyen HB, Gupta G. Comput Biol Chem.2021 Jun; 92: 107475. doi: 10.1016/j. compbiolchem.2021.107475. Epub 2021 Mar 17. PMID: 33813188.

[12] Nguyen M, Brettin T, Long SW, Musser JM, Olsen RJ, Olson R, Shukla M, Stevens RL, Xia F, Yoo H, Davis JJ.2018. Developing an in silico minimum inhibitory concentration panel test for *Klebsiella pneumoniae*. *Sci Rep* 8: 421.

[13] Camp P - JV, Haslam DB, Porollo A.2020. Prediction of antimicrobial resistance in Gram - negative bacteria from whole - genome sequencing data. *Front Microbiol* 11: 1013.

[14] Moradigaravand D, Palm M, Farewell A, Mustonen V, Warringer J, Parts L.2018. Prediction of antibiotic resistance in *Escherichia coli* from large - scale pan - genome data. *PLoS Comput Biol* 14: e1006258.

[15] Her HL, Wu YW.2018. A pan - genome - based machine learning approach for predicting ntimicrobial resistance activities of the *Escherichia coli* strains. *Bioinformatics* 34: i89–i95

[16] Mao J, Guan S, Chen Y, Zeb A, Sun Q, Lu R, Dong J, Wang J, Cao D., Application of a deep generative model produces novel and diverse functional peptides against microbial resistance. *Comput Struct Biotechnol J*.2022 Dec 19; 21: 463 - 471. doi: 10.1016/j. csbj.2022.12.029. eCollection 2023. PMID: 36618982