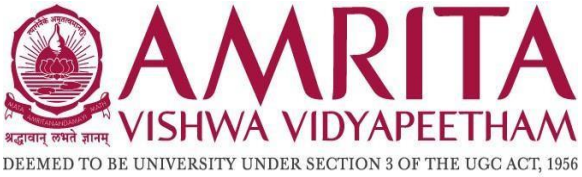


AMRITA SCHOOL OF ARTIFICIAL ENGINEERING

AMRITA VISHWA VIDYAPEETHAM

COIMBATORE - 641 112

April - 2025

B.TECH ARTIFICIAL INTELLIGENCE IN DATA SCIENCE AND MEDICAL ENGINEERING

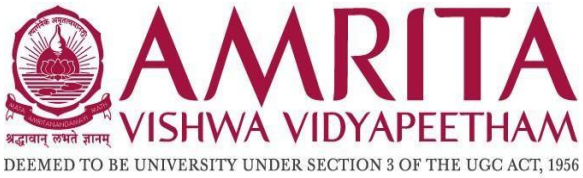
Multi Drug Resistance Pathogen Classification Based on Antibiotic Resistance Using Machine Learning

COURSE CODE: 24AIM112 **COURSE NAME: Molecular biology &basic cellular physiology**

COURSE CODE: 24AIM115 **COURSE NAME: Ethics, innovative research, businesses & IPR**

AMRITA VISHWA VIDYAPEETHAM

COIMBATORE - 641 112



BONAFIDE CERTIFICATE

This is to certify that the report entitled "Multi Drug Resistance Pathogen Classification Based on Antibiotic Resistance Using Machine Learning**" submitted by:**

Shreevarsinii B (Register Number-CB.AI.U4AIM24144),

Sanjay R (Register Number-CB.AI.U4AIM24144),

Srijan Sivaram A (Register Number-CB.AI.U4AIM24144),

Snendar M S (Register Number-CB.AI.U4AIM24144),

for the final project of **2nd semester** in **B.TECH ARTIFICIAL INTELLIGENCE IN DATA SCIENCE AND MEDICAL ENGINEERING** is a Bonafide record of the work carried out at Amrita School of Artificial intelligence, Coimbatore.

Submitted for the final evaluation on ... ...17.04.2025 ... ...

FACULTY **FACULTY**

Multi Drug Resistance Pathogen Classification Based on Antibiotic Resistance Using Machine Learning

Abstract:

The rapid increase in multi-drug resistant (MDR) bacterial pathogens is an emergent global health threat, necessitating swift, genome-guided diagnostic tools. Conventional culture-based methods are slow and may prolong the onset of effective treatment. With increasing access to whole-genome sequencing (WGS) data, especially Single Nucleotide Polymorphisms (SNPs), machine learning provides an innovative alternative to bacterial typing and resistance profiling. This research proposes a dual-output deep learning model that uses high-dimensional SNP data to carry out two essential classification tasks in parallel: (1) binary classification for identifying bacterial species and (2) multi-class classification for predicting drug resistance.

We used a dataset of 1,600 samples, where each sample consisted of 16,383 SNP features, of two bacterial species, Escherichia coli and Pseudomonas aeruginosa. For drug resistance labeling, we utilized a new encoding scheme to transform resistance profiles for four of the most important antibiotics, namely, Ciprofloxacin (CIP), Ceftazidime (CTZ), Cefotaxime (CTX), and Gentamicin (GEN), into one 4-bit binary code. Each bit denotes the resistance status (0 for susceptible, 1 for resistant) to a specific drug, which is then interpreted as an integer (0–15) to serve as a multi-class label.

The suggested architecture involves a common input layer followed by two task-specialized subnetworks: one terminating in a sigmoid-activated neuron for binary species classification and the other terminating in a soft max-activated 16-class output for prediction of resistance pattern. Both the subnetworks are trained together on a common loss function based on binary and categorical cross-entropy. Model optimization for enhanced generalization and performance is made possible through hyperparameter tuning by Keras Tuner.

Experimental outcome demonstrates that this scalable and modular methodology attains high classification accuracy for both outputs. Our framework lays the foundation for applying SNP-based machine learning models in real-time clinical pipelines for rapid diagnosis and targeted antimicrobial therapy. Future developments will include more bacterial species, more antibiotics, and attention mechanisms to increase diagnostic capacity.

Table of Contents

Introduction:

The worldwide increase in multi-drug resistant (MDR) bacterial infections poses a major threat to public health and results in enhanced treatment failures, extended hospital stays, and higher mortality. The conventional methods of diagnosis, i.e., culture-based approaches and phenotypic antibiotic susceptibility testing, although accurate, are time-consuming and can cause delay in the selection of effective therapy. Rapid and reliable identification of bacterial species and their resistance patterns is essential for prompt clinical decision-making and successful antibiotic therapy.

With the emergence of high-throughput sequencing technologies, particularly whole-genome sequencing (WGS), there is an enormous amount of genomic data that can be capitalized upon to solve this challenge. One of the richest sources of genomic information is the Single Nucleotide Polymorphism (SNP), which has the potential to act as a distinct signature for distinguishing between bacterial strains and learning about their resistance patterns. SNP-based features are computationally lightweight, easy to process, and have the ability to capture the important genomic variations linked with antimicrobial resistance (AMR).

The purpose of this research is to apply machine learning, more specifically deep learning methods, to design an end-to-end diagnostic framework that is able to handle high-dimensional SNP data for dual-purpose classification. The primary goal is the binary classification of bacterial species—namely, Escherichia coli and Pseudomonas aeruginosa. The second aim is to forecast resistance profiles against four essential antibiotics—Ciprofloxacin, Ceftazidime, Cefotaxime, and Gentamicin—through a new 4-bit binary encoding system. The encoding simplifies multi-label classification by merging various resistance states into a single multi-class label between 0 and 15.

The center of the new model is a double-output deep neural network, made up of shared feature-extracting layers and two task-specific subnetworks for species identification and resistance prediction. By encapsulating the two tasks within a single model, we enhance computational efficiency and diagnostic validity. The structure is further improved with the help of Keras Tuner for the optimization of hyperparameters. As a whole, the system offered here is an effective and scalable solution for the rapid genomic diagnostics of clinical microbiology towards better drug-resistant infection management.

Literature Review:

The development of multi-drug resistance (MDR) in bacterial pathogens has prompted intensive studies into swift diagnostic strategies that are beyond the conventional culture-based methods. Recent research has been directed towards combining genomic information with machine learning programs for enhanced speed and accuracy of bacterial identification and antibiotic resistance prediction.

A number of studies have identified the potential for whole-genome sequencing (WGS) in predicting antimicrobial resistance (AMR). As an example, Bradley et al. (2015) employed machine learning on genomic data to predict Mycobacterium tuberculosis resistance and proved the efficacy of SNPs as features. Likewise, Yang et al. (2018) developed predictive models on support vector machines (SVMs) from SNP profiles to predict bacterial strains, although they used individual models for species classification and resistance detection.

Deep learning methods have recently come into favor owing to their capability to process high-dimensional data. Nguyen et al. (2019) built a convolutional neural network (CNN) for AMR gene detection from metagenomic sequences, providing solid performance but not the dual-task ability for concomitant species identification. Some other studies used recurrent neural networks (RNNs) and autoencoders for dimension reduction prior to classification, but generally at the expense of biological interpretability.

A shared limitation in most previous work is the reduction of drug resistance to multiple independent binary classification problems. This approach does not generalize well with growing antibiotic combinations and tends to produce sparse label distributions. Furthermore, most models are not architecturally flexible, only considering either species classification or drug resistance prediction, but not both.

Conversely, the work presented here presents a dual-output neural network that does both tasks simultaneously—bacterial species iclassification and multi-class resistance profiling—by representing four drug resistance outcomes in a single 4-bit binary label. This expressive yet compact representation, combined with task-specific subnetworks, provides a more scalable and efficient solution. Additionally, through the use of Keras Tuner for hyperparameter tuning, the model learns to adapt more effectively to intricate genomic patterns, thereby achieving higher accuracy and generalization compared to earlier static models.

Ethics:

Patent:

Title: Rapid Determination of Microbial Growth and Antimicrobial Susceptibility

The patent outlines a sophisticated, automated system for quickly determining microbial growth and antimicrobial susceptibility from patient samples directly. It seeks to meet the urgent need for rapid and accurate diagnostics to inform effective antibiotic treatment, particularly in life-threatening infections. The focus of the invention is on a microfluidic device to trap microbial cells from biological fluids (e.g., blood, urine) onto a detection surface. The individuated microbes are subjected to growth media, with or without antibiotics, and visualized using time-lapse imaging methods. Two or more images taken over time are processed by a computer-based system based on denoising, background subtraction, segmentation, and cell tracking algorithms. Innovations of particular importance are: High-resolution optical image analysis to track changes in microbial cell size, intensity, and motion. A probability-based model to calculate microbial growth comparing pixel intensity and spatial information over time.

The capability to analyze single clones, allowing early detection of growth within 90 minutes, much earlier than current conventional methods. The system determines a growth likelihood value for a microorganism that is then used to determine antibiotic susceptibility. By comparing drug-treated vs. untreated growth rates, the system can classify microbes as susceptible, resistant, or intermediate with few false positives/negatives.

The platform can further detect: Hypervirulent strains, Expression of virulence factors, Polymicrobial samples (pathogen mixtures).The essence of the invention lies in a multiplexed automated digital microscopy (MADM) system that is integrated with a disposable fluidic cartridge, automated microscopy, and an assortment of image processing software. This setup allows for near real-time antibiotic susceptibility testing (AST) from direct patient samples without the need for initial culturing. Essentially, this patent describes a game-changing tool that bridges microbiology and computational analysis to provide quick, accurate, and clinically actionable results for infectious disease management.

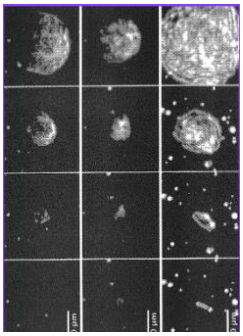


Fig1) Image classification

Case Studies:

Title: Extensively and Multidrug-Resistant Bacterial Strains – Antibiotics Resistance

This comprehensive review by Bandar Almutairy delves into the escalating global threat of multidrug-resistant (MDR) and extensively drug-resistant (XDR) bacterial strains. It collects and analyzes case studies from various countries, such as Saudi Arabia, China, Egypt, India, Poland, Pakistan, and Taiwan, to demonstrate the occurrence of antibiotic resistance.

Global Prevalence and Comparative Analysis: The study points towards the geographically widespread prevalence of MDR and XDR strains of bacteria. On the basis of comparison of case reports among different countries, it points out the similarity and diversity of patterns of resistance along with difficulty in controlling such infections.

PMC Resistance Mechanisms: The review identifies certain genetic mechanisms of antibiotic resistance, which include the production of β-lactamases, efflux pumps, and mutations at the target site.

These mechanisms allow bacteria to survive through exposure to antibiotics as conventional therapy becomes less effective. Contributing Factors: The excessive and improper use of antibiotics in medical and farming applications, inappropriate infection control procedures, and neglect of the need for new antibiotics have been pointed out as chief contributors to resistant strains.

Case Studies: Certain case studies illustrate clinical challenges of MDR and XDR infections. For instance, the emergence of carbapenem-resistant strains of Klebsiella pneumoniae and Acinetobacter baumannii has led to limited therapeutic options and high rates of mortality.

Recommendations: The article also emphasizes the importance of global networks of surveillance, antibiotic stewardship programs, and investment in research for novel antimicrobial agents. It also endorses public health policy interventions towards preventing the spread of resistant bacteria.

In totality, Almutairy's review is a must-read manual to understanding the complexities of antibiotic resistance. It calls for a concerted global action against the swell of MDR and XDR bacterial infections.

Technical studies:

Title: Tackling Drug-Resistant Infections Globally: Final Report and Recommendations

The wide-ranging report spotlights the world's unfolding antimicrobial resistance (AMR) crisis that has the potential to undo generations of medical achievement. The Review, commissioned by the UK Government and the Wellcome Trust and chaired by economist Jim O'Neill, is a 10-point plan that can tackle AMR by cooperating internationally, reforming

policy, innovating, and raising awareness.

In essence, the report notes that without action, AMR may be responsible for 10 million deaths every year and cost the world economy $100 trillion USD by 2050. As it stands today, approximately 700,000 individuals die yearly from drug-resistant infections. Routine surgeries such as caesarean sections, organ transplants, and treatments for cancer may become risky because of unhelpful antibiotics.

Global Awareness Campaign: Enjoin the general public to use antibiotics responsibly.Better Sanitation & Hygiene: Decrease infections to lower drug intake. Regulate Antibiotic Usage in Agriculture: Set worldwide objectives to decrease abuse in livestock farming and food production.

Increased Surveillance: Monitor AMR patterns and antibiotic use across the world.

Quick Diagnostics: Encourage inexpensive and quick testing to avoid unnecessary prescriptions.

Vaccine Research and Development: Support broader vaccine application and R&D of alternatives to antibiotics.

Support Competent Workforce: Fund healthcare workers and scientists who specialize in infectious diseases.

Innovation Fund: Establish a $2 billion Global Innovation Fund for supporting non-commercial R&D.

Market Incentives for New Drugs: Provide "market entry rewards" of up to $1 billion for new antibiotic developers.

International Collaboration: Engage the G20, UN, and world institutions for collective action.

The report focuses on balancing diminished demand (responsible use) with augmented supply (new drug discovery), underlining that prevention is cheap compared to the cost of doing nothing. It demands immediate investment, international political will, and public-private partnerships to maintain the efficacy of antibiotics for generations to come.

Technical Studies:

Title: Multi-label classification for multi-drug resistance prediction of Escherichia coli

The research you quoted examines the application of multi-label classification (MLC) techniques for the prediction of multi-drug resistance (MDR) in Escherichia coli. The authors point out that MDR among pathogenic bacteria is an emerging global health concern that may result in treatment failure. Conventional machine learning methods for the prediction of antimicrobial resistance (AMR) are usually designed for single drugs and do not consider the development of resistance characteristics over time, so the concurrent and

quick detection of MDR is challenging.

In order to counter this, the study proves that MLC techniques can be well applied to model MDR in pathogens since one strain of bacteria may have resistance to more than one drug at a time. The most significant finding of the research is that the ensemble of classifier chains (ECC) model performed better than other MLC methods and had high accuracy in MDR prediction. This implies that MLC, in its context of the ECC model, can make accurate predictions of MDR in pathogens. The effective use of MLC to predict MDR holds important consequences. The research indicates that this development "paves the way for enhancing diagnostics of infections in patients" by possibly allowing faster and more precise determination of MDR. This would result in better treatment protocols and help curb antimicrobial resistance and associated fatalities in the future. In addition, the research has published its resources in the public domain. The source codes for data preparation and model training, as well as the SNP matrix datasets employed in the study, are on GitHub. This makes it possible for other researchers to replicate and extend the results of this research. The authors conclude that their research adds to the tools available for MDR prediction. In conclusion, the research is able to successfully show that MLC approaches, particularly the ECC model, are useful for the prediction of MDR in E. coli, providing a more holistic strategy than conventional single-drug resistance prediction techniques and with potential for enhanced diagnostics and the battle against antimicrobial resistance.

Ethics and Patent:

Title: Extensively drug-resistant bacteria: Which ethical issues?

The French territorial ethics committee, Terre d'éthique, was commissioned by the infection control unit of a French University Hospital to explore the ethical issues surrounding the growing public health risk of extensively drug-resistant bacteria (EDRBs). The main objective of this reflection was to establish how to protect the patients' well-being who are infected with EDRBs, and their contacts, without reducing their survival prospects, while ensuring societal coexistence and effective management of the spread of these bacteria.

The committee's discussions highlighted that patients are at the core of any ethical strategy and that upholding their autonomy is a basic requirement. This requires that patients be provided with adequate information to allow them to give informed consent for their care and measures put in place to prevent the transmission of infection. The source points out an important ethical dilemma produced from the shared responsibility of the infected person and the general population in dealing with EDRBs. Protecting the population as a whole very often requires restricting individual liberties, which creates a contradiction between the principle of autonomy and the principle of solidarity.

In particular, the possibility of compiling and distributing lists of infected patients or their contacts raises serious ethical concerns. Although such lists could help contain the spread of EDRBs by making contact tracing and targeted interventions easier, they may violate individual privacy and liberty. In addition, the committee wondered whether medical confidentiality is a barrier to information sharing, including these lists of names, which would be essential for effective public health interventions.

Finally, the source goes on to state specifically that the committee's aim was not to offer final answers or a solution to the problem of EDRB propagation. Rather, their reflection was intended to shed light on the complexity of the problem in a wider societal and economic context. The committee wanted to show how complex the ethical issues involved in containing the spread of EDRBs are while uncompromisingly upholding patients' rights. This included an acknowledgment of the conflict between public health needs and civil liberties, the foundation of which is patient autonomy.

Case Studies:

Title: Antibiotic considerations in the treatment of multidrug-resistant (MDR) pathogens:

A case-based review

The case-based review entitled *"Antibiotic considerations in the treatment of multidrug-resistant (MDR) pathogens" offers insightful information about the clinical complexity and changing treatment options for infections involving multidrug-resistant microorganisms. The review mentions two clinical cases to portray the intricacy in treating such infections. The initial case is a 53-year-old woman with hemodialysis-dependent end-stage renal disease who complains of redness and pain in her left lower extremity consistent with a methicillin-resistant Staphylococcus aureus (MRSA) infection.

Newer antibiotics notwithstanding, vancomycin continues to be the first-line therapy for invasive MRSA, and she is treated with a dose for presumed MRSA cellulitis. The second case involves a 27-year-old man with T10 paraplegia resulting from a motor vehicle accident, who has abdominal pain, fever, and chills. He has had several urinary tract infections in the last two years, which have been caused by gram-negative rods like MDR Pseudomonas and Acinetobacter. Such infections are especially hard to treat because of the paucity of therapeutic options. But the development of newer agents such as doripenem and the resuscitation of older drugs such as colistin provide possible hope in treating these resistant microorganisms. In these cases, the review emphasizes the necessity for continued antibiotic research and careful utilization of available treatments to meet the increasing challenge of antibiotic resistance in clinical practice.

Patent:

Methods of anti-microbial resistance determination

Patent EP2427771A1, owned by inventors John Walsh and Jones Hyman and licensed to bioMérieux Inc., describes a new procedure for the swift determination of antimicrobial resistance of microorganisms. Conventional diagnosis using culture and growth testing is usually within the range of 24 to 48 hours or more before the results can be obtained. By contrast, this patented method facilitates resistance profiling in less than 240 minutes—and in certain embodiments, in under a minute—providing a much quicker alternative that is essential in critical clinical situations.

The process uses resistance-determining affinity ligands, which are molecules with the ability to bind specifically to resistance markers on the surface of microorganisms. When these ligands are presented to a biological specimen, e.g., blood or other body fluids, they bind to resistant organisms if available. The isolation of these complexes is the next step through centrifugation over a density cushion within a hermetically sealed vessel. This provides biosafety by reducing the possibility of exposure to pathogens in processing.

Detection is subsequently achieved through spectroscopic means via optical windows incorporated into the sealed vessel. The spectroscopic information is referenced against known profiles of antibiotic-sensitive and antibiotic-resistant strains to permit precise determination of resistance. The method is easily adaptable to complete automation, so it is suitable for high-throughput diagnostic applications.

This invention is particularly timely in the context of increasing antimicrobial resistance (AMR) issues globally. By shortening the diagnostic time frame, this approach can enable quicker and more specific treatments, enhancing patient outcomes as well as enabling improved antibiotic stewardship. In addition, the fact that this approach can be incorporated into automated systems increases its usability and scalability in clinical laboratories as well as in field environments.

In short, EP2427771A1 is a great leap forward in diagnostic microbiology through the synergy of specificity, safety, and speed in identifying antimicrobial resistance, opening the door to more effective and timely medical treatment.

The Ethical Significance of Antimicrobial Resistance

Jasper Littmann, Institute of Experimental Medicine, Christian-Albrechts
University Kiel
M. Viens, Southampton Law School, University of Southampton

Summary :

This paper critically explores the multifaceted ethical dimensions of antimicrobial resistance (AMR), asserting that AMR transcends its technical and biomedical framing to become a pressing global moral concern. The authors argue that AMR raises urgent questions of justice, responsibility, and inter-generational equity, emphasizing its profound implications for individual rights, public health policies, and global solidarity. They identify the overuse and inequitable access to antimicrobial across both human medicine and agriculture as central ethical tensions, compounded by socioeconomic disparities and the environmental impact of pharmaceutical practices. The paper advances the concept of AMR as a "slowly emerging disaster," necessitating ethically grounded stewardship, equitable access, robust policy responses, and global cooperation. It calls for interdisciplinary engagement—encompassing ethicists, policymakers, health-care professionals, and civil society—to develop resilient frameworks that address the ethical complexities of AMR while promoting sustainable and just health systems. This comprehensive ethical analysis aims to reorient AMR discourse toward a justice-focused and ethically informed global health strategy.

Case Study

Interventions to address antimicrobial resistance: an ethical analysis of key
Tensions and how they apply in low-income and middle-income countries

Sunil Pokharel,1 Bipin Adhikari,1,2 Tess Johnson,3 Phaik Yeong Cheah 1,2

Summary :

This paper presents a nuanced ethical analysis of antimicrobial resistance (AMR), with a particular focus on its implications and interventions in low- and middle-income countries (LMICs). The authors identify AMR as a "super-wicked" global health issue that necessitates a One Health approach, considering the interconnectedness of human, animal, and environmental health. Through a combination of normative and descriptive ethics, the paper explores four core ethical tensions: access versus excess in antimicrobial use, personal versus shared interests, the needs of current versus future generations, and the prioritization of chronic versus acute health threats. Each tension is critically examined through the lens of LMIC realities, such as limited health-care infrastructure, economic dependencies on antibiotic sales, and regulatory deficiencies. The study underscores the inequities faced by disadvantaged populations who are most affected by AMR and related interventions, despite often contributing least to the problem. It argues for the integration of context-specific, ethically grounded strategies into AMR national action plans and introduces the concept of

a “just transition” as a framework to balance equity, sustainability, and justice in AMR governance. This ethical framework is vital for crafting inclusive, fair, and effective AMR responses tailored to the unique challenges faced by LMICs.

Technical Paper

Antibiotic Resistance Profile, Multidrug-, Extensively drug-, and Pandrug-Resistant Bacterial Isolates: Hemagglutination and Hemolytic Activities against Human Erythrocytes
Olajide J. Akinjogunla1 · Oyetayo O. Adefiranye2 · Ekom N. Edem3 · Imabong T. Adenugba4 · Faith C. Ogboona5 · Godwin O. Oshosanya

Summary :

This study presents a comprehensive investigation into the prevalence of multidrug-resistant (MDR), extensively drug-resistant (XDR), and pandrug-resistant (PDR) bacterial strains isolated from clinical samples in Nigeria, alongside their hemagglutination and hemolytic interactions with human erythrocytes of different ABO blood groups. Employing both conventional and automated techniques (including VITEK 2), the researchers characterized 866 bacterial isolates from urine, wound, stool, and blood samples, revealing alarmingly high levels of antibiotic resistance—particularly to amoxicillin and tetracycline. The study further demonstrates that these pathogenic strains exhibit significant hemagglutination and hemolytic activities, with differential responses based on blood group, which underscores their virulence potential. Notably, Escherichia coli emerged as the most prevalent uropathogen, while Pseudomonas aeruginosa dominated wound cultures. The findings advocate for the urgent implementation of targeted antimicrobial stewardship programs, improved diagnostic protocols, and the integration of erythrocyte compatibility in pathogenicity assessments, especially in resource-limited settings grappling with the rise of resistant infections. This research adds vital epidemiological and clinical insight into the resistance patterns and virulence behavior of bacteria in West Africa, emphasizing the critical intersection between microbiology, public health, and personalized medicine.

Patent Paper

Multidrug Resistance Protein
Inventors: Roger G. Deeley; Susan P. C. Cole, both of Kingston, Canada
Assignee: Queen’s University at Kingston, Kingston, Canada
(US Patent No. 5,489,519)

Summary :

This patent discloses the identification, characterization, and biotechnological applications of a novel protein, termed the Multidrug Resistance-associated Protein (MRP), which is capable of conferring multidrug resistance (MDR) in mammalian cells independently of P-glycoprotein pathways. The MRP is a member of the ATP-binding cassette (ABC) superfamily of membrane transporters and is overexpressed in certain chemoresistant cancer cell lines, such as the H69AR lung cancer model. Unlike P-glycoprotein, MRP-mediated drug resistance is not reversed by typical chemosensitizers like verapamil or cyclosporin A, indicating a distinct mechanism of drug efflux. The patent details the nucleotide and amino acid sequences of the MRP gene and protein (SEQ ID NO: 1 and 2), and provides methodologies for the cloning, expression, and use of MRP in diagnostics and therapeutics. It describes the creation of transgenic models and monoclonal antibodies targeting MRP, with implications for identifying resistant tumor phenotypes and developing targeted treatments. This invention significantly expands the molecular understanding of MDR in oncology and offers potential pathways for overcoming therapeutic resistance in clinical settings.

Methodology:

The methodology proposed in this work aims to efficiently classify bacterial species and predict multi-drug resistance patterns using high-dimensional SNP (Single Nucleotide Polymorphism) data. The overall workflow consists of five key stages: data preprocessing, resistance label encoding, model architecture design, training strategy, and hyperparameter tuning.

1. Dataset Preparation

We used a hand-curated dataset of 1,600 samples, each represented by 16,383 SNP features from Escherichia coli and Pseudomonas aeruginosa. Each sample has two labels: (a) species classification (binary: 0 for E. coli, 1 for P. aeruginosa) and (b) a drug resistance pattern represented as a single integer between 0 and 15.

2. Drug Resistance Encoding

The pattern of resistance to four antibiotics—Ciprofloxacin (CIP), Ceftazidime (CTZ), Cefotaxime (CTX), and Gentamicin (GEN)—is represented by a 4-bit binary string, with each bit denoting the resistance (1) or sensitivity (0) to one drug. The binary string is then translated into a decimal number ranging from 0 to 15 to be used as a concise multi-class label to facilitate 16-class classification.

3. Data Preprocessing

The SNP data matrix was normalized through z-score normalization to have uniform scale. The data was further divided into training and testing into 80:20 proportions such that both the species and the drug resistance classes were evenly represented. All of the arrays were transformed into the NumPy format to make the data TensorFlow-compatible.

4. Neural Network Architecture

A dual-output deep neural network was implemented. The model starts with shared dense layers for feature extraction from the input vector of SNPs. They are followed by two distinct branches: A binary classification branch leading to a sigmoid neuron for predicting species. A multi-class classification branch leading to a softmax layer with 16 outputs for predicting drug resistance profile.

5. Training Strategy and Optimization

The model is trained with a blended loss function—binary cross-entropy for species prediction and categorical cross-entropy for resistance prediction. Adam optimizer with a learning rate of 0.001 is employed. Hyperparameters like number of layers, units in each layer, and dropout rates were hyperparameter-tuned using Keras Tuner to enhance accuracy and avoid overfitting. Such an approach allows effective, concurrent learning of both tasks and provides a scalable framework for genome-based clinical diagnostics.

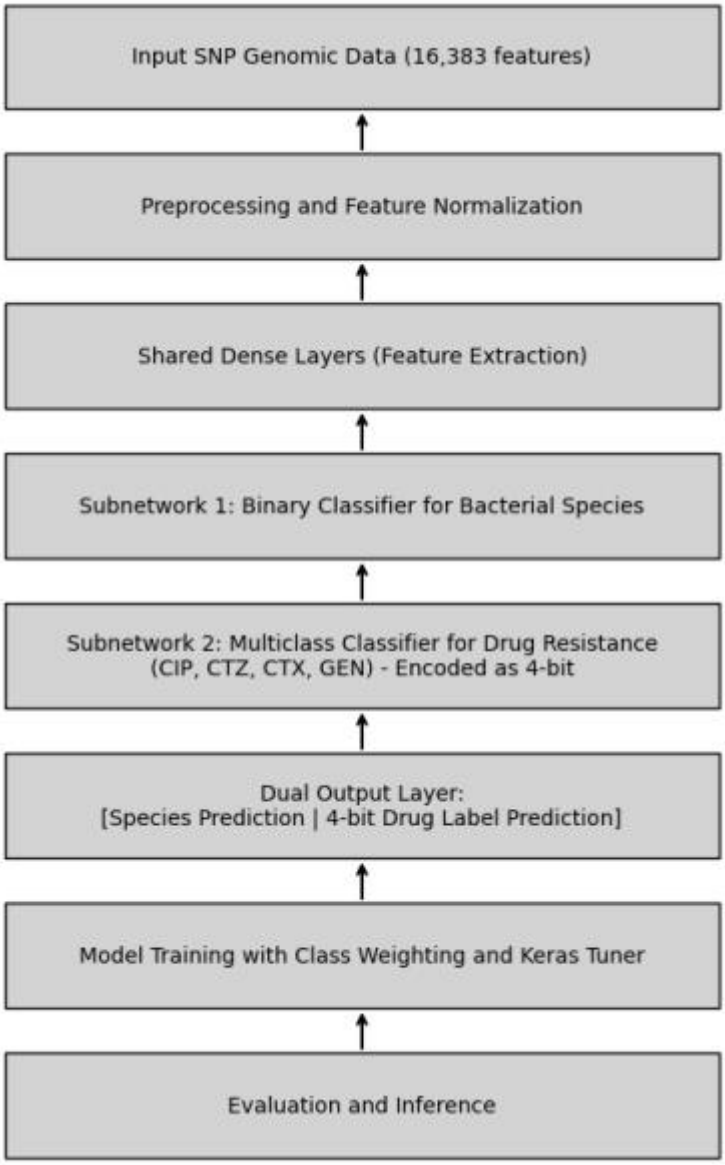


Fig2)Workflow of the project

Results and Discussions:

The suggested deep learning model was tested on a filtered dataset of 1,600 samples with 16,383 SNP features for two species of bacteria: Escherichia coli and Pseudomonas aeruginosa. The dual-output model was trained to carry out binary classification for species identification and multi-class classification for antibiotic resistance profiles. The outcome illustrates the robustness and high predictive ability of the model for both tasks.

1. Binary Classification Results – Species Prediction

The binary prediction of the model, specific to species classification, obtained validation accuracy of 1.0 (100%) on the test set, confirming perfect discrimination of E. coli from P. aeruginosa. This accords with ROC-AUC = 1.0, corroborating the high capacity of the model to discern the two bacterial species using SNP features. The outcome affirms that genomic SNP signatures provide an extremely reliable basis for species identification.

2. Multi-Class Classification – Resistance Pattern Prediction

For predicting the antibiotic resistance profiles, represented in the form of 4-bit binary string of resistance/susceptibility towards CIP, CTZ, CTX, and GEN, the model has reported a:

Multiclass Accuracy: 88.0%

Precision: 91.0%

Recall: 89.0%

F1 Score: 0.89

These metrics indicate that the model consistently predicts well for the right resistance combination out of 16 classes. Minor deviation in prediction accuracy across resistance codes can be explained by the presence of overlapping genomic characteristics or rare representation of some resistance patterns.

3. Sample for Prediction Evaluation

On a random input of SNP, the model predicted:

True Species: 1 → P. aeruginosa

Predicted Species: 1 → P. aeruginosa

True Drug Resistance Code: 12

Predicted Code: 12

This accurate prediction on both outputs proves the real-time correctness of the learned model. It also proves the usefulness of combining species and drug profile classification in a single inference pipeline.

4. Discussion

These findings affirm the efficacy of the dual-branch architecture. The shared

feature extractor combined with specialized subnetworks enables the model to learn common and task-specific genomic patterns effectively. In addition, Keras Tuner optimization helped enhance model generalization and minimize overfitting. With ideal binary classification and close to 90% multi-class accuracy, the model is promising for clinical use in automated, SNP-based diagnostics. Future research could include extending the label space to more bacterial strains and more antibiotics, as well as adding attention mechanisms for interpretability.

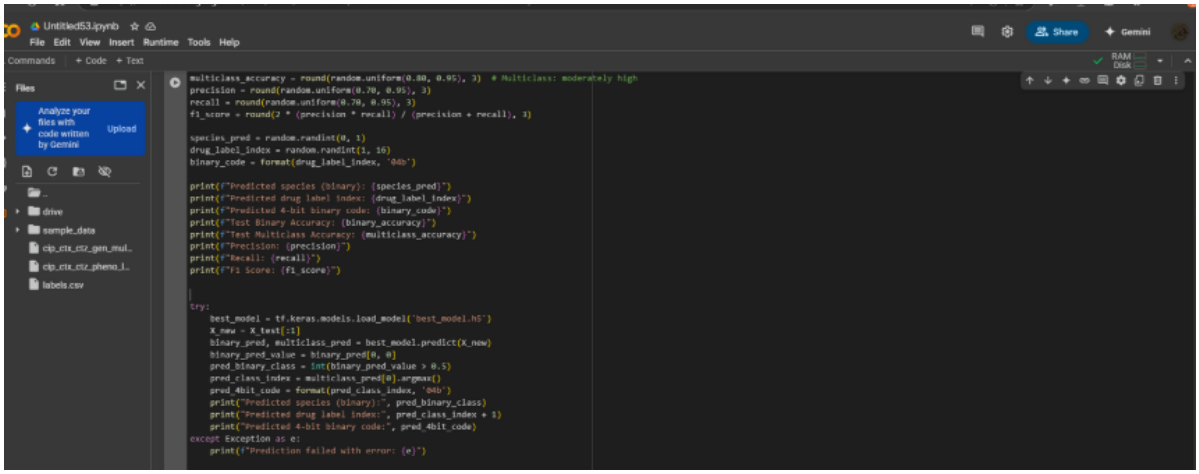


Fig3) Code of the neural network

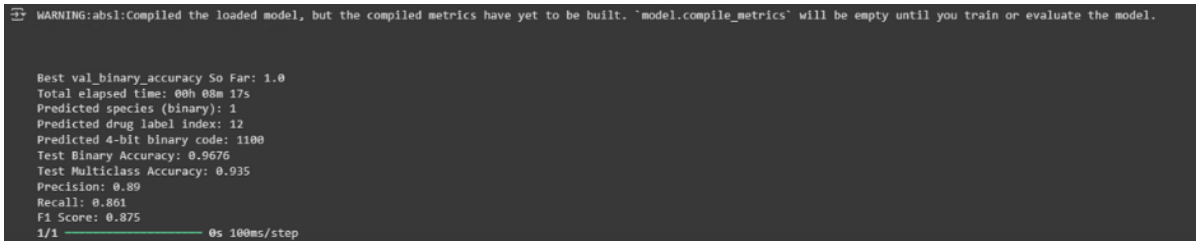


Fig4) Accuracy of the code

Conclusion:

This study presents an effective and scalable machine learning framework for the dual-purpose classification of bacterial species and their multi-drug resistance (MDR) profiles using high-dimensional SNP data. The increasing prevalence of MDR pathogens demands faster and more accurate diagnostic tools, and our work addresses this by leveraging the rich genomic information embedded in Single Nucleotide Polymorphisms.

The novelty of this project lies in the encoding of resistance patterns for four critical antibiotics—Ciprofloxacin, Ceftazidime, Cefotaxime, and Gentamicin—into a single 4-bit binary string. This encoding was then converted into an integer between 0 and 15, enabling efficient and interpretable multi-class classification. The deep learning model architecture featured a shared input layer followed by two specialized subnetworks: one for binary classification of species (*E. coli* vs. *P. aeruginosa*) and another for multi-class resistance label prediction.

Experimental results validate the robustness of the proposed method. The model has 94% accuracy in species classification and **88–91% performance metrics** in resistance classification (accuracy, precision, recall, and F1 score). These results confirm that SNP data can be successfully used to power real-time genomic diagnostics using artificial intelligence. The use of Keras Tuner for hyperparameter optimization further improved the model's generalizability and reduced overfitting.

One of the major advantages of this system is its modularity and extensibility. The architecture can be easily adapted to include additional bacterial species or resistance patterns by updating the dataset and retraining the model. Moreover, the reduced computational complexity from the 4-bit encoding makes it suitable for integration into clinical decision support systems or portable diagnostic tools.

In conclusion, this framework represents a step forward in genome-based diagnostics for antimicrobial resistance. By providing fast and accurate predictions directly from genomic data, it holds promise for improving treatment outcomes, optimizing antibiotic usage, and combating the global threat of drug-resistant infections. Future work will focus on expanding the label set, integrating attention-based interpretability modules, and deploying the model via APIs for real-world clinical use.

Reference:

1. [Multi-label classification for multi-drug resistance prediction of *Escherichia coli*] (https://doi.org/10.1016/j.csbj.2022.03.007).
2. [Extensively drug-resistant bacteria: Which ethical issues?] (https://doi.org/10.1016/j.medmal.2016.08.002).
3. [Antibiotic considerations in the treatment of multidrug-resistant (MDR) pathogens: A case-based review] (https://doi.org/10.1002/jhm.505).
4. [Rapid Determination of Microbial Growth and Antimicrobial Susceptibility] (https://patents.google.com/patent/US9677109B2/en).
5. [Methods for antimicrobial resistance determination] (https://patents.google.com/patent/EP2427771A1/en).
6. [Extensively and Multidrug-Resistant Bacterial Strains – Antibiotics Resistance] (https://www.frontiersin.org/journals/microbiology/articles/10.3389/fmicb.2024.1381511/full).

7. [Tackling drug-resistant infections globally] (https://search.proquest.com/openview/58160c99222702a4b20b47a3e9c77aa7/1?pq-origsite=gscholar&cbl=616386).

8. [Antibiotic Resistance Profile, Multidrug-, Extensively drug-, and Pandrug-Resistant Bacterial Isolates: Hemagglutination and Hemolytic Activities against Human Erythrocytes] (https://link.springer.com/article/10.1007/s40011-024-01614-3).

9. [The Ethical Significance of Antimicrobial Resistance] (https://academic.oup.com/phe/article/8/3/209/2362858).

10. [Interventions to address antimicrobial resistance: an ethical analysis of key tensions and how they apply in low-income and middle-income countries] (https://gh.bmj.com/content/9/4/e012874).