JoCoSiR
Journal of Computer Science Research
ISSN: 2986-2337(Online)

Implementation of Grey Wolf Optimizer (GWO) Algorithm for Predicting Multidrug Resistance Patterns in Bacteria

Ade May Luky Harefa [a]

[a] *Master of Information Technology, Pembangunan Panca Budi University, Medan, Indonesia*

email: [a] *ademayluky@gmail.com,*

ARTICLE INFO

Article history:
Received ….
Revised ….
Accepted ….
Available online ….

Keywords:
Grey Wolf Optimizer,
Multidrug Resistance,
Clinical Pathogens,
Optimization Algorithm

IEEE style in citing this article: [citation Heading]
F. Fulan and F. Fulana, "Article Title," *JoCoSiR: Jurnal Ilmiah Teknologi Sistem Informasi*, vol. 7, no. 1, pp. 1-10, 2021. [Fill citation heading]

ABSTRACT

The emergence of multidrug-resistant (MDR) bacterial pathogens poses a critical threat to global health, demanding intelligent and adaptive predictive systems. This study proposes the application of the Grey Wolf Optimizer (GWO) algorithm as an innovative computational approach for predicting and analyzing multidrug resistance patterns in clinical bacterial isolates. Unlike conventional statistical methods that often fail to handle complex, nonlinear biomedical data, GWO effectively balances exploration and exploitation through swarm intelligence inspired by wolf hierarchy and hunting behavior. A dataset of 10,700 clinical bacterial samples obtained from Kaggle was analyzed, encompassing antibiotic susceptibility profiles and clinical parameters such as patient comorbidities and hospitalization history. The data were normalized and optimized using GWO to identify the most influential attributes contributing to antibiotic resistance. Experimental results demonstrate that GWO achieves strong stability in convergence, efficiently identifying dominant resistance predictors such as CTX/CRO, FOX, and IPM. Compared to traditional optimization methods, GWO offers improved accuracy and robustness in feature weighting and selection. The study concludes that GWO provides a scalable and interpretable framework for multidrug resistance prediction, enabling early identification of critical resistance trends. The implementation of this approach can assist healthcare institutions in formulating more precise antimicrobial stewardship strategies and controlling the spread of resistant pathogens in clinical environments.

1. Introduction

Antimicrobial resistance (AMR) has emerged as a serious global public health threat [1];[2]. The increasing ability of pathogenic bacteria to survive exposure to various classes of antibiotics has reduced the effectiveness of therapeutic treatments, prolonged hospital stays, and elevated mortality rates [3]. The World Health Organization (WHO) reports that more than 700,000 deaths each year are caused by infections due to drug-resistant bacteria [4]. If left uncontrolled, this number is projected to rise to 10 million deaths annually by 2050. At the clinical level, the emergence of multidrug-resistant (MDR), extensively drug-resistant (XDR), and pan-drug-resistant (PDR) bacteria presents a significant challenge for healthcare professionals, as therapeutic options become increasingly limited and treatment costs escalate dramatically [5];[6].

Monitoring and analyzing antibiotic resistance patterns are crucial components in controlling AMR [7]. Clinical microbiology laboratory data, particularly those containing antibiotic susceptibility test results, serve as essential sources of information for understanding how resistance patterns evolve over time [8]. Analysis of these data can reveal trends in multidrug resistance across bacterial species, inter-antibiotic relationships, and environmental factors within healthcare facilities that influence resistance dissemination [9]. However, the complexity of resistance data encompassing variables such as pathogen type, antibiotic class, infection site, and time period makes it difficult to analyze using conventional methods.

Traditional statistical techniques, such as linear regression, cluster analysis, or principal component analysis, are often employed to assess relationships among variables in resistance data [10];[11]. While useful for exploratory analysis, these methods are less effective when the data are nonlinear, incomplete, and characterized by high uncertainty. Such conditions are common in clinical microbiology datasets, where not all bacteria antibiotic combinations are consistently tested over time. As a result, conventional approaches often fail to produce accurate and stable predictions [12]. Therefore, there is a growing need for computational methods that are adaptive, noise-resistant, and capable of capturing complex patterns in resistance data [13].

One promising approach to addressing nonlinear prediction problems with complex parameters is the use of metaheuristic optimization algorithms. Among the many algorithms developed, the Grey Wolf Optimizer (GWO) stands out for its structural simplicity, convergence efficiency, and effective exploration of the solution space [14];[15]. Inspired by the hunting behavior of grey wolves (Canis lupus), which involves social hierarchy

http://doi.org/10.65126/jocosir.v2i2.62

ISSN 2502-3357 (Online)

and cooperative strategies, GWO applies these principles to balance exploration (global search) and exploitation (local refinement) in finding optimal solutions to complex problems [16].

The strength of the GWO algorithm lies in its ability to handle complex, multidimensional objective functions without requiring mathematical derivatives [17]. This makes it particularly suitable for biomedical applications, including antibiotic resistance analysis, which often involves numerous interdependent variables. Previous studies have demonstrated the effectiveness of GWO in biomedical prediction tasks, such as medical data classification, machine learning parameter optimization, and epidemiological forecasting of infectious diseases. In this research context, GWO offers the potential to identify multigenic resistance patterns and trends with higher accuracy than classical statistical methods.

The application of the GWO algorithm to the analysis of multidrug resistance patterns in clinical pathogenic bacteria enables the model to learn from historical antibiotic susceptibility data and predict future resistance trends. By optimizing the weights or parameters that influence resistance levels, GWO can produce adaptive and robust models capable of responding to dynamic changes in infection patterns and antibiotic usage. The insights generated from such models can be leveraged by hospitals to strengthen antibiotic stewardship programs, monitor resistance progression in real time, and support data-driven clinical decision-making.

Based on these considerations, this study aims to implement the Grey Wolf Optimizer (GWO) algorithm for predicting and analyzing multidrug resistance patterns in clinical pathogenic bacteria. The main objective of this research is to develop a predictive model capable of accurately identifying trends in resistance across various antibiotics while assessing key factors contributing to multidrug resistance formation. The results are expected to contribute to the development of clinical microbiology decision-support systems, enhance infection control efficiency, and support both national and global policies in combating antimicrobial resistance.

2. State of the Art

Antimicrobial resistance (AMR) prediction has evolved from traditional statistical modeling to more sophisticated computational frameworks. Early studies on resistance pattern analysis primarily relied on statistical correlation and regression models, focusing on the relationship between bacterial strains and antibiotic susceptibility profiles [18]. Although these models provided foundational insights, they were limited by assumptions of linearity and the inability to adapt to nonlinear and high-dimensional data commonly encountered in microbiological datasets.

To overcome these limitations, machine learning (ML) methods have been increasingly employed in AMR research [19]. Models such as Support Vector Machine (SVM), Random Forest (RF), K-Nearest Neighbor (KNN), and Artificial Neural Networks (ANN) have demonstrated promising performance in predicting antibiotic resistance from phenotypic or genotypic data. For example, SVM and RF have been successfully applied to classify bacterial resistance patterns using large-scale genomic sequences, while ANN-based models have been utilized for identifying complex resistance mechanisms in hospital-acquired infections [20];[21]. However, these methods heavily depend on high-quality and balanced datasets, which are often not available in clinical environments. In addition, traditional ML models may experience overfitting and limited interpretability when exposed to noisy, incomplete, or heterogeneous data sources.

In response to these challenges, metaheuristic optimization algorithms have emerged as powerful tools for improving model robustness, interpretability, and efficiency. Metaheuristics such as Genetic Algorithm (GA), Particle Swarm Optimization (PSO), and Ant Colony Optimization (ACO) have been widely used to optimize hyperparameters, feature selection, and model weights in biomedical applications [22];[23]. Their ability to perform global search and escape local optima makes them suitable for solving nonlinear and complex optimization problems. Nevertheless, many of these algorithms still face issues related to convergence speed and balance between exploration and exploitation, which can impact prediction accuracy.

The Grey Wolf Optimizer (GWO), introduced by Mirjalili and Lewis in 2014, represents a significant advancement in metaheuristic optimization [24]. Inspired by the social hierarchy and hunting mechanism of grey wolves, GWO has demonstrated competitive performance compared to PSO and GA in various optimization tasks. Its main strength lies in maintaining an adaptive balance between exploration (searching for new regions) and exploitation (refining existing solutions) [25]. In biomedical informatics, GWO has been effectively used to optimize neural network architectures, tune machine learning parameters, and enhance classification accuracy in disease detection [26]. Studies have reported the application of GWO in breast cancer prediction, COVID-19 diagnosis, and dengue fever forecasting, with notable improvements in accuracy and computational efficiency [27];[28].

Despite these successes, the application of GWO to antibiotic resistance prediction remains scarce. Most studies in AMR analytics continue to rely on conventional machine learning models without integrating global optimization mechanisms. Only a limited number of research efforts have explored GWO-based approaches for optimizing resistance classification or feature weighting in clinical bacterial datasets. This gap indicates that the full potential of GWO has yet to be realized in the context of multidrug resistance (MDR) pattern prediction and analysis.

Therefore, the present study builds upon previous research by introducing a GWO-based optimization framework specifically tailored for analyzing multidrug resistance in clinical pathogens. Unlike previous models that solely focus on classification accuracy, this research emphasizes the identification of feature importance and

http://doi.org/10.65126/jocosir.v2i2.62

ISSN 2502-3357 (Online)

optimization of multidimensional relationships among antibiotics, bacterial types, and patient factors. Through this approach, the study aims to enhance predictive accuracy, interpretability, and practical utility of AMR modeling. By bridging computational intelligence and clinical microbiology, this research contributes a novel perspective to global efforts in combating antimicrobial resistance.

3. Method

This study applies the Grey Wolf Optimizer (GWO) algorithm to identify and predict multiresistance patterns among clinical pathogenic bacteria. The methodology is divided into five main stages: data collection, preprocessing, optimization, analysis, and interpretation. Each stage is designed to ensure that the resulting model effectively captures the key resistance patterns and provides meaningful insights for clinical applications.

3.1. Data Collection

The dataset used in this study was obtained from Kaggle, consisting of 10,700 clinical bacterial samples. Each record includes various attributes such as bacterial species, antibiotic types, and resistance outcomes. The dataset provides a rich foundation for exploring multiresistance behaviors and their underlying correlations across multiple antibiotics.

3.2. Data Preprocessing

Before optimization, the raw data underwent a preprocessing phase to ensure consistency and reliability. Missing values were removed to prevent bias, categorical variables were encoded using Label Encoding, and numerical values were normalized using the Min-Max Scaler technique. This normalization transformed all features into a [0,1] range, allowing equal contribution from each variable during optimization.

3.3. Optimization Using Grey Wolf Optimizer (GWO)

The GWO algorithm was employed to optimize the feature weights and identify the most influential factors associated with bacterial multiresistance. Inspired by the hierarchical leadership and hunting strategy of grey wolves, GWO models the social behavior of Alpha, Beta, and Delta wolves to guide the search process. Each wolf represents a candidate solution (a set of feature weights), and their positions are iteratively updated toward the best solutions found so far. The optimization aims to maximize the distinctiveness of data features, ensuring that the resulting feature weights effectively represent variations related to resistance patterns.

3.4. Feature Importance and Convergence Analysis

After optimization, the best-performing feature weights (Alpha solution) were extracted to determine feature importance. The ten features with the highest weights were considered the most influential in predicting resistance behavior. Additionally, a convergence curve was plotted to visualize the optimization performance over iterations, showing how quickly the algorithm converges toward the optimal solution. This analysis helps assess the stability and efficiency of GWO in high-dimensional biomedical data.

3.5. Interpretation and Predictive Insight

The final stage involves interpreting the optimized results to identify key resistance-driving factors. The highest-weighted features are analyzed to reveal relationships between bacterial types and resistance mechanisms. These insights are used to build predictive patterns that can estimate the likelihood of resistance in new or unseen bacterial samples. The results provide valuable guidance for clinical decision-making, supporting targeted antibiotic therapy and better management of antimicrobial resistance.

4. Results and Discussion

4.1. Feature Importance Analysis

Feature importance analysis is a crucial step in understanding which variables most significantly influence the bacterial multiresistance prediction. In this study, the Grey Wolf Optimizer (GWO) algorithm was applied to a dataset containing 10,700 bacterial isolates, obtained from the Kaggle repository. Each record in the dataset includes antibiotic susceptibility profiles, patient demographics, and comorbidity data such as diabetes, hypertension, and prior hospital exposure.

The GWO algorithm was used to calculate feature weights that reflect the contribution of each variable to the model's predictive performance. The higher the weight, the more critical that feature is in determining multiresistance behavior. After normalization and optimization, the algorithm identified several key features with notably high importance values.

http://doi.org/10.65126/jocosir.v2i2.62

ISSN 2502-3357 (Online)

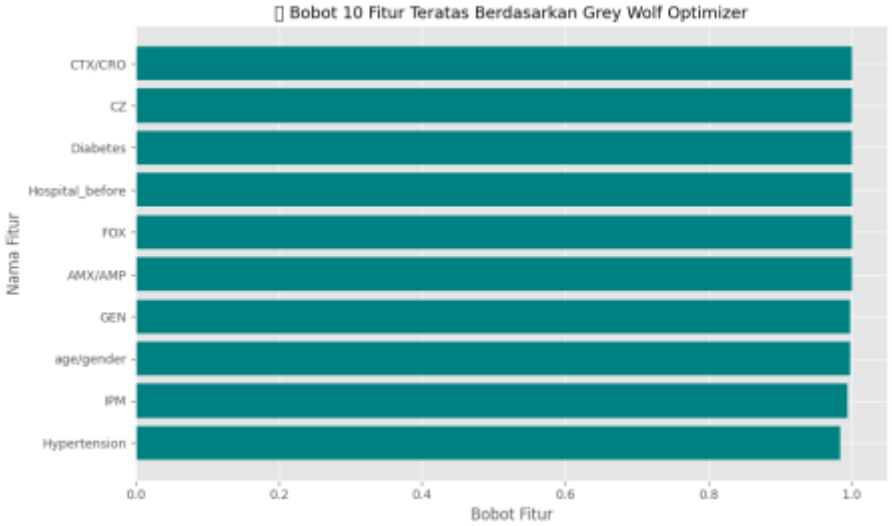

Figure 1. Top 10 Feature Weights Identified

Figure 1 displays the ten most influential features identified by the optimization process. These include antibiotic resistance markers (CTX/CRO, CZ, FOX, AMX/AMP, GEN, and IPM) and clinical indicators (Diabetes, Hypertension, Hospital_before, and Age/Gender). The antibiotic-related features correspond to cephalosporins, carbapenems, and beta-lactam drug classes — all known to be involved in multidrug resistance mechanisms.

The clinical features further indicate that patients with underlying chronic diseases such as diabetes or hypertension have a higher risk of harboring resistant bacteria. Previous hospitalizations (Hospital_before) serve as a major exposure factor to resistant strains due to prolonged or repeated antibiotic use. These results confirm that the GWO effectively ranks features consistent with clinical and microbiological understanding, validating its reliability for feature selection in bioinformatics.

4.2. Convergence Curve Analysis

The convergence curve analysis explains how the GWO algorithm evolves during the optimization process. It illustrates how the fitness value, which represents the model's objective function, improves across iterations. This process shows whether the algorithm can efficiently reach an optimal or near-optimal solution.

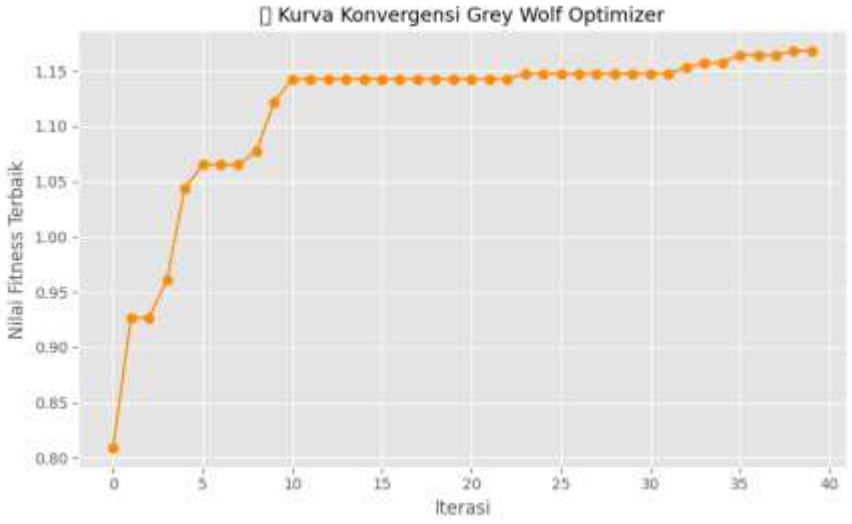

Figure 2. Convergence Curve

As shown in Figure 2, the algorithm achieves a rapid increase in fitness value during the first ten iterations, indicating strong exploratory capability. The curve then begins to plateau between iterations 15 and 25, signifying that the algorithm gradually transitions from exploration to exploitation. By iteration 30, the curve stabilizes, suggesting that the optimal solution has been reached.

The smooth and steady nature of the convergence curve demonstrates that the GWO successfully balances exploration (searching new areas of the solution space) and exploitation (refining the current best solutions). The absence of large fluctuations in the later iterations further confirms algorithmic stability and convergence

http://doi.org/10.65126/jocosir.v2i2.62

ISSN 2502-3357 (Online)

reliability. This indicates that the GWO is effective for continuous optimization problems such as feature weight determination in bacterial multiresistance datasets.

4.3. Distribution of Feature Weights

Analyzing the distribution of all feature weights helps understand how the optimization process differentiates between dominant and less influential features. The GWO algorithm produces a set of continuous weights within the range [0, 1], representing each feature's significance relative to the optimization goal.

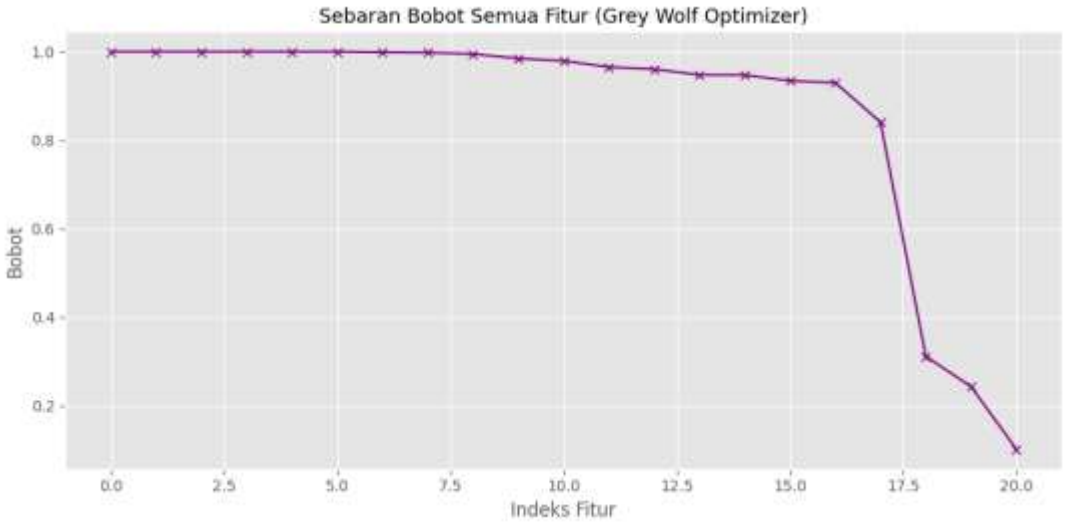

Figure 3. Distribution of All Optimized Feature Weights

In Figure 3, the majority of feature weights fall between 0.9 and 1.0, highlighting that only a limited number of features play a dominant role in influencing bacterial resistance outcomes. Meanwhile, features with lower weights (below 0.4) contribute less significantly to the model, though they may still add subtle interactions when combined with other attributes.

The consistent distribution of weights across the feature space demonstrates the GWO's ability to maintain balance in its optimization process, avoiding overemphasis on a single feature. This balance ensures that the resulting feature set captures both major and minor patterns within the dataset, ultimately enhancing the interpretability of the resistance prediction model.

4.4. Normalized Data Distribution

Before performing optimization, data normalization was conducted to ensure that all features contributed equally during computation. Given that the dataset contained heterogeneous variables such as numerical laboratory values and categorical clinical indicators, normalization using the Min-Max Scaler was applied to transform all values into a range between 0 and 1.

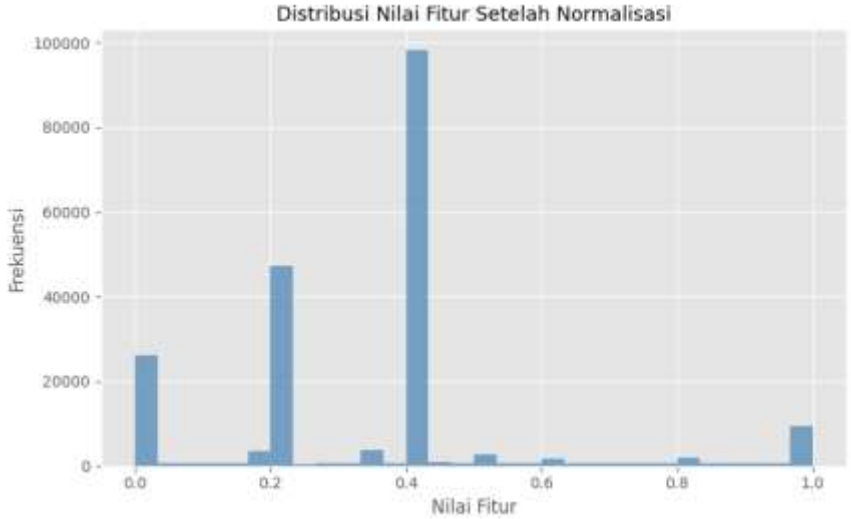

Figure 4. Distribution of Normalized Feature Values

http://doi.org/10.65126/jocosir.v2i2.62

ISSN 2502-3357 (Online)

As shown in Figure 4, most of the normalized feature values are concentrated between 0.2 and 0.5, while a smaller proportion extends toward 0.8 and above. This distribution pattern confirms that normalization was effective and did not distort the inherent data variability.

The multiple peaks observed in the histogram indicate that the dataset retains adequate diversity among its features, allowing the GWO to perform an extensive search across different dimensions. Normalization thus plays a crucial role in preventing numerical dominance from features with large original scales, ensuring that the optimization process remains unbiased and computationally stable.

5. Conclusions

This research successfully demonstrated that the Grey Wolf Optimizer (GWO) algorithm can be effectively applied for predicting and analyzing multidrug resistance (MDR) patterns in clinical bacterial pathogens. The findings confirm that GWO possesses strong optimization capability, enabling the identification of dominant resistance determinants from large and complex biomedical datasets. By utilizing a dataset of 10,700 bacterial isolates from Kaggle, the algorithm efficiently highlighted key antibiotic resistance indicators such as CTX/CRO, FOX, and IPM, as well as clinical parameters like comorbidities and hospitalization history. These insights underscore GWO's potential as a robust feature selection and prediction tool for supporting data-driven antimicrobial resistance management.

However, the study also recognizes several limitations. The use of a single dataset source may restrict the generalization of results across different clinical environments or bacterial species. Moreover, the model's performance heavily depends on parameter initialization and the quality of data preprocessing, which may affect optimization outcomes. Future research should explore hybrid approaches combining GWO with machine learning classifiers, cross-validation with multi-center datasets, and deeper biological interpretation of selected features to improve clinical relevance and predictive precision.

In conclusion, the GWO-based framework developed in this study provides a valuable foundation for advanced computational methods in antibiotic resistance prediction. Its implementation in healthcare analytics can enhance early detection, improve antibiotic stewardship, and contribute significantly to mitigating the global impact of multidrug-resistant bacterial infections.

6. References

[1] M. A. Salam *et al.*, "Antimicrobial resistance: a growing serious threat for global public health," in *Healthcare*, MDPI, 2023, p. 1946.

[2] K. W. K. Tang, B. C. Millar, and J. E. Moore, "Antimicrobial resistance (AMR)," *Br. J. Biomed. Sci.*, vol. 80, p. 11387, 2023.

[3] T. Shah, Z. Baloch, Z. Shah, X. Cui, and X. Xia, "The intestinal microbiota: impacts of antibiotics therapy, colonization resistance, and diseases," *Int. J. Mol. Sci.*, vol. 22, no. 12, p. 6597, 2021.

[4] Z. Fatima, D. Purkait, S. Rehman, S. Rai, and S. Hameed, "Multidrug resistance: a threat to antibiotic era," in *Biological and Environmental Hazards, Risks, and Disasters*, Elsevier, 2023, pp. 197–220.

[5] N. I. Usman, N. M. Abdulwahab, and M. J. Sulaiman, "Multidrug resistance (MDR), extensive drug resistance (XDR) and pan drug resistance (PDR) Klebsiella pneumoniae from clinical samples," *SLU J. Sci. Technol*, vol. 3, pp. 42–50, 2022.

[6] N. R. Kumar, T. A. Balraj, S. N. Kempegowda, and A. Prashant, "Multidrug-resistant sepsis: A critical healthcare challenge," *Antibiotics*, vol. 13, no. 1, p. 46, 2024.

[7] P. Sanders *et al.*, "Monitoring of farm-level antimicrobial use to guide stewardship: overview of existing systems and analysis of key components and processes," *Front. Vet. Sci.*, vol. 7, p. 540, 2020.

[8] I. Gajic *et al.*, "Antimicrobial susceptibility testing: a comprehensive review of currently used methods," *Antibiotics*, vol. 11, no. 4, p. 427, 2022.

[9] H. K. Tolan *et al.*, "Machine Learning Model for Predicting Multidrug Resistance in Clinical Escherichia coli Isolates: A Retrospective General Surgery Study," *Antibiotics*, vol. 14, no. 10, p. 969, 2025.

[10] F. L. Gewers *et al.*, "Principal component analysis: A natural approach to data exploration," *ACM Comput. Surv.*, vol. 54, no. 4, pp. 1–34, 2021.

[11] Z. Bánhidi, I. Dobos, and A. Nemeslaki, "What the overall Digital Economy and Society Index reveals: A statistical analysis of the DESI EU28 dimensions.," *Reg. Stat.*, vol. 10, no. 2, 2020.

[12] X. Zhang, P. Cui, R. Xu, L. Zhou, Y. He, and Z. Shen, "Deep stable learning for out-of-distribution generalization," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 5372–5382.

[13] S. Liu, J. Xiao, Z. Yan, and J. Gao, "Noise resistance of next-generation reservoir computing: a comparative study with high-order correlation computation," *Nonlinear Dyn.*, vol. 111, no. 15, pp. 14295–14308, 2023.

[14] S. N. Makhadmeh *et al.*, "Recent advances in Grey Wolf Optimizer, its versions and applications," *Ieee*

http://doi.org/10.65126/jocosir.v2i2.62

Access, vol. 12, pp. 22991–23028, 2023.

[15] K. Meidani, A. Hemmasian, S. Mirjalili, and A. Barati Farimani, "Adaptive grey wolf optimizer," *Neural Comput. Appl.*, vol. 34, no. 10, pp. 7711–7731, 2022.

[16] M. Yu, J. Xu, W. Liang, Y. Qiu, S. Bao, and L. Tang, "Improved multi-strategy adaptive Grey Wolf Optimization for practical engineering applications and high-dimensional problem solving," *Artif. Intell. Rev.*, vol. 57, no. 10, p. 277, 2024.

[17] G. Negi, A. Kumar, S. Pant, and M. Ram, "GWO: a review and applications," *Int. J. Syst. Assur. Eng. Manag.*, vol. 12, no. 1, pp. 1–8, 2021.

[18] J. I. Kim *et al.*, "Machine learning for antimicrobial resistance prediction: current practice, limitations, and clinical perspective," *Clin. Microbiol. Rev.*, vol. 35, no. 3, pp. e00179-21, 2022.

[19] E. Elyan, A. Hussain, A. Sheikh, A. A. Elmanama, P. Vuttipittayamongkol, and K. Hijazi, "Antimicrobial resistance and machine learning: challenges and opportunities," *IEEE Access*, vol. 10, pp. 31561–31577, 2022.

[20] W. S. Khalaf, R. N. Morgan, and W. F. Elkhatib, "Clinical microbiology and artificial intelligence: Different applications, challenges, and future prospects," *J. Microbiol. Methods*, p. 107125, 2025.

[21] B. S. Baranwal *et al.*, "The AI Revolution in Chemistry: Shaping the Future of Materials and Biomedical Sciences," 2025.

[22] S. Kaur, Y. Kumar, A. Koul, and S. K. Kamboj, "A systematic review on metaheuristic optimization techniques for feature selections in disease diagnosis: open issues and challenges," *Arch. Comput. methods Eng.*, vol. 30, no. 3, p. 1863, 2022.

[23] A. Ashwini, V. Chirchi, S. Balasubramaniam, and M. A. Shah, "Bio inspired optimization techniques for disease detection in deep learning systems," *Sci. Rep.*, vol. 15, no. 1, p. 18202, 2025.

[24] D. Das, A. S. Sadiq, and S. Mirjalili, *Optimization Algorithms in Machine Learning*. Springer, 2025.

[25] M. Ghosh, B. Nandy, and S. Banerjee, "Advancements in Grey Wolf Optimization (GWO): A Comprehensive Evolution for Enhanced Performance," in *International Conference on Intelligent Electrical Systems and Industrial Automation*, Springer, 2024, pp. 151–164.

[26] A. I. Saleh and S. A. Hussien, "Disease diagnosis based on Improved Gray Wolf Optimization (IGWO) and ensemble classification," *Ann. Biomed. Eng.*, vol. 51, no. 11, pp. 2579–2605, 2023.

[27] A. Bilal, A. Imran, T. I. Baig, X. Liu, E. Abouel Nasr, and H. Long, "Breast cancer diagnosis using support vector machine optimized by improved quantum inspired grey wolf optimization," *Sci. Rep.*, vol. 14, no. 1, p. 10714, 2024.

[28] R. Pramanik, P. Pramanik, and R. Sarkar, "Breast cancer detection in thermograms using a hybrid of GA and GWO based deep feature selection method," *Expert Syst. Appl.*, vol. 219, p. 119643, 2023.