International Scientific Journal of Engineering and Management
ISSN: 2583-6129
Volume: 03 Issue: 04 | April – 2024
DOI: 10.55041/ISJEM01501
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

ANTIBIOTICS RESISTANCE PREDICTION USING MACHINE LEARNING ALGORITHMS

Bharathi A
Assistant Professor of
Information Technology
Nandha Engineering College
Erode
Bharathi.a@nandhaengg.org

Gokul A
Information Technology
Nandha Engineering College
Erode
gokulannadurai2002@gmail.com

Naveen Kumar T
Information Technology
Nandha Engineering College
Erode
naveenkumarthangaraj2002@gmail.com

Ranjith V
Information Technolgy
Nandha Engineering College
Erode
ranjithtsv@gmail.com

ABSTRACT

Opioid Risk Antibiotics Resistance abuse, overdose, and drug addiction have become public health concerns as a result of the notable increase in the use of both prescribed and over-the-counter medications in the United States. The growing incidence of opioid usage presents a complex public health issue, one of which is the possible influence on antibiotic resistance. The existing system predicts antibiotic resistance profiles in bacterial infections among hospitalized patients using machine learning (ML) algorithms and electronic medical records (EMRs). However, it ignores the impact of opioid use, a major public health issue that may be connected to antibiotic resistance. The complex association between opioid-related behaviors and antibiotic resistance may not be adequately captured by powerful machine learning algorithms, despite their use. The larger context of antibiotic resistance patterns is limited by the lack of linkage with data on opioid intake. The proposed abstract aims to improve predictive accuracy and understanding of the complex factors influencing antibiotic resistance emergence by integrating support vector machines (SVM) and linear regression to thoroughly analyse the relationship between opioid risk and antibiotic resistance. This approach addresses these limitations. We suggested combining linear regression and support vector machines (SVM) to better understand the complex connection between opioid risk and antibiotic resistance. In order to pinpoint the main causes of increased risk, trends in opioid consumption data are analyzed using the SVM model. In parallel, the relationship between opioid abuse and the rise in antibiotic resistance is examined using linear regression. Our goal is to have a thorough grasp of the complex relationship between opioid-related behaviors and the emergence of antibiotic resistance by combining these two analytical approaches. This project is designed under netbeans with java as front end. Weka tool for machine learning analysis

Keywords: Antibiotic Resistance, Machine Learning, Bacterial Infections, Global Crisis, Healthcare.

1. INTRODUCTION

Antibiotic resistance has become a major danger to public health in recent years, casting doubt on the effectiveness of once-true therapies for bacterial diseases. Conventional means of overcoming this resistance have not worked well, thus new strategies are needed to deal with this urgent problem [1]. Machine learning algorithms have shown great promise in the battle against antibiotic resistance due to their capacity to examine large datasets and spot complex patterns [2]. Researchers can better understand the intricate mechanisms driving antibiotic resistance and create more focused and efficient antibiotic usage methods by utilizing artificial intelligence [3]. Modern healthcare and technology together have the power to completely transform how we respond to antibiotic resistance and provide promise for long-term solutions that will protect the future of world health [4].

The increasing and difficult problem of antibiotic resistance in the field of global health seriously jeopardizes our ability to successfully tackle bacterial illnesses [5]. Data on participants' health problems, particular athletic activities, and sociodemographic traits were gathered via a questionnaire [6]. This happens when bacteria adapt and create defenses against the effects of antibiotics, making once-effective drugs less effective or even nonexistent [7]. Antibiotic abuse and misuse in agriculture, healthcare, and other fields accelerates the emergence of this resistance, raising serious concerns for public health [8]. Because of the intense competition in this sympatric lifestyle, organisms have devised strategies to survive, including the manufacture of antibiotic compounds and/or antibiotic resistance genes [9]. The risk of incurable infections increases as our supply of antibiotics declines, endangering the results of standard medical operations and weakening the advancements accomplished in contemporary medicine. Antibiotic resistance must be addressed immediately, necessitating a broad and coordinated effort to create novel approaches, encourage appropriate antibiotic usage, and guarantee the ongoing efficacy of these life-saving drugs [10].

Machine learning, a revolutionary branch of artificial intelligence that allows computers to learn from data and become more efficient over time, has completely changed how we tackle complicated issues and make predictions. The human microbiomes of various bodily locations, including the gut, airways, skin, urine, and genitalia, have been the subject of a multinational endeavor for more than ten years [11]. Fundamentally, machine learning imitates the cognitive processes of human learning by enabling computers to see patterns, derive conclusions, and make judgments without the need for explicit programming. This paradigm change has broad ramifications for a number of areas, including technology, healthcare, and finance. Machine learning allows for the iterative learning of large datasets by algorithms, which may be used to find hidden connections, automate operations, and improve decision-making [12]. As a constantly developing discipline, machine learning keeps pushing the envelope of what is conceivable, providing creative answers to complex problems and opening the door for a day when intelligent systems will be crucial in determining the direction of technology.

2. RELATED WORKS

Mascaro Valentina, [13] et al. has suggested in this research that athletes who play collision or contact sports are more likely to become infected with Staphylococcus aureus. Frequent use of antibiotics, close contact circumstances, skin injuries, sharing of sports equipment, and poor hygiene are risk factors for colonization and infections. Consequently, our particular objectives were to ascertain the incidence of S. antimicrobial resistance of S. aureus among various contact sport players, which were little examined in the European region. A cross-sectional investigation was carried out among Italian athletes who compete in contact or collision sports to determine the carriage of aureus isolates and its correlation with specific sport-related risk variables. Data on participants' health problems, particular athletic activities, and sociodemographic traits were gathered via a questionnaire. Tests were conducted to determine the antibiotic susceptibility and isolation of nares, oropharynx, and finger swabs. According to our findings, there is a significant resistance to the antimicrobials used to treat infections as well as a high prevalence of carriage. The goal of prevention measures was to lower S. Enhancing aureus colonization might potentially decrease infections. An opportunistic human pathogen known as Staphylococcus aureus affects people of all ages and genders. It can cause skin and soft tissue infections (SSTIs) as well as more serious illnesses including bloodstream infections or pneumonia. has grown resistant to a number of medications, and methicillin-resistant (MRSA) infections are becoming more prevalent in the community as well as among hospitalized patients.

Baron, Sophie A., [13] et al. Human microbiomes, as suggested in this research, are intricate ecosystems made up of bacteria, viruses, archaea, and/or eukaryotes that are coevolving in an environment under a variety of selective pressures, including the use of antibiotics, nutrition, and/or lifestyle choices. Because of the intense competition in this sympatric lifestyle, organisms have devised strategies to survive, including the manufacture of antibiotic compounds and/or antibiotic resistance genes (ARGs). This environment becomes a major source of antibiotic-resistant genes (ARGs) for harmful bacteria, increasing the risk of infection from these germs. The study of these microbiomes has led to the development of two complimentary techniques, culture and metagenomics, which have been used to better understand the types of bacteria and ARGs found in the human body as well as the variables influencing their diversity and abundance. The goal of this review was to define resistance genes and multidrug-resistant bacteria that have already been isolated from the human microbiota using culturomics and metagenomics techniques, as well as to identify factors that influence the colonization and propagation of ARGs and/or multidrug-resistant bacteria. Joshua Lederberg originally described the microbiome as the biological community of harmful, commensal, and symbiotic bacteria that coexist in human bodies. The human microbiomes of various bodily locations, including the gut, airways, skin, urine, and genitalia, have been the subject of a multinational endeavor for more than ten years. Two main approaches used in this research are metagenomics and culturomics. New microbial species have been found as a result of this effort, which has also made it possible to identify certain microbiota and site characteristics. Antibiotic resistance genes (ARGs) found in the microbiome have been investigated among them.

Somme, Morten O.A.,[14] et al. has suggested in this work that our capacity to cure infectious diseases is being jeopardized by human pathogenic bacteria's rising levels of multi-drug resistance. The study of antibiotic resistance reservoirs that pathogens may access is becoming more and more popular because antibiotic resistance determinants are easily transferred across bacteria through lateral gene transfer. Owing to the increased probability of interaction and genetic interchange with pathogens during the course of an illness, particular consideration should be given to the human microbiota, which may be the most easily accessible source of resistance genes. In fact, a great deal of antibiotic resistance has been shown in cultured isolates from the human microbiota in earlier research. We recently showed that the functional repertoire of resistance genes in the human microbiome is far more varied than previously anticipated using previous culture-dependent approaches by employing metagenomics functional selections. We demonstrated that a large number of resistance genes from human fecal sample cultivated proteobacteria are similar to resistance genes carried by human pathogens, offering compelling evidence for the recent genetic exchange of this resistance mechanism. By comparison with all known genes in public databases, however, the majority of resistance genes that we found using culture independent metagenomics sampling from the same samples were unique.

Pouwels, Koen B,[15] et al. has suggested in this work that there is a direct correlation between the rates of bacterial resistance and antibiotic usage. Research examining the relationship between antibiotic use and resistance at the population level typically concentrate on the tenuous relationships that exist between the use of a particular antibiotic or group of antibiotics and resistance against that particular antibiotic.2,4,5 The influence of using other antibiotics and the fact that (multiple) resistance genes are frequently linked, allowing co-selection, are generally disregarded. If there is a decrease in one antibiotic and a rise in another (related) drug, this might be seen as confounding by other antibiotics. For instance, considering the almost total cross-resistance between these two medicines, replacing amoxicillin with ampicillin would not be expected to result in a change in amoxicillin resistance. Furthermore, by eliminating rival bacterial flora, therapy with one drug may raise the density of organism’s resistant to another agent within a patient.6 Co-selections can happen when resistance genes are connected to the same mobile genetic element, which functions as a resistance "vector." For instance, it is commonly known that these elements frequently connect the ampicillin and trimethoprim resistance genes. Consequently, while prescribing ampicillin, one looks for genes associated to trimethoprim resistance in addition to amoxicillin resistance.

Costelloe. Powell [16] et al. has suggested in this work to do a thorough literature review and, where necessary, to conduct meta-analyses of research looking at the development of antibiotic resistance in people who get primary care antibiotic prescriptions. We have conducted a systematic review and meta-analysis of studies where the impact of antimicrobial use on the emergence of resistance has been assessed for individual patients in primary care because, to the best of our knowledge, no systematic reviews have been published in this area [17], Only a small number of high-quality studies have reported on the relationship between prescribing and prevalence of resistance for individuals treated in primary care. Not only did we want to determine which medicines were most and least likely to produce resistance, but we also wanted to quantify the degree and duration of any link. Two reviewers (CC and one other) independently assessed the quality of the entire articles and extracted data using a specially created form. Disputes were settled by conversation with an additional author (ADH). The paper's author was contacted for clarification in cases where data extraction proved challenging or ambiguous.

In this research, Ce´ire Costelloe [18] et al. have suggested conducting a systematic assessment of the literature and, if necessary, conducting meta-analyses of studies that look into the development of antibiotic resistance in people who get antibiotic prescriptions in primary care. The rising incidence of antibiotic resistance is one of the most urgent issues facing healthcare systems. Increased by fewer new drugs being introduced into clinical settings, this resistance is generally acknowledged as posing a serious risk to public health.1. In general practice, there are worries that illnesses caused by bacteria resistant to antibiotics may take longer to heal and that certain common infections are getting harder to cure.2. Use of antibiotics improperly or indiscriminately can lead to some antimicrobial resistance. As a result, efforts are being made at the municipal, state, and federal levels to encourage "antibiotic stewardship," with the aim of enhancing the appropriateness of the use of antibiotics. However, the effectiveness of these programs depends on prescribers and patients receiving ongoing education, which must be backed by solid research connecting the use of antibiotics and the evolution of resistance. Primary care continues to provide the bulk of antibiotics to patients, despite the fact that several nations have been successful in lowering the amount of antibiotics recommended.3. 4 A significant portion of this use is for the treatment of suspected respiratory infections, and levels of prescription vary greatly both within and between nations, indicating the possibility of additional reductions.

In this research, Stefan [19] Wager et al. have proposed Treatment effect heterogeneity is a key concept in many scientific and engineering problems, from customized marketing suggestions to individualized medicine. In this study, we expand Breiman's popular random forest approach to create a non-parametric causal forest for predicting diverse treatment effects. We demonstrate that causal forests have an asymptotically Gaussian and centered sampling distribution and are pointwise consistent for the true treatment effect in the framework with confoundedness. We also provide a useful technique for building asymptotic confidence intervals centered at the causal forest estimates for the true treatment impact. Our theoretical findings are predicated on a broad family of random forest algorithms' generic Gaussian theory. This is the first collection of results that we are aware of that permits the use of any kind of random forest, including regression and classification forests, for statistical inference that can be proven to be valid. In our testing, we discover that causal forests are significantly more effective than traditional approaches that rely on nearest-neighbor matching, particularly when there are unimportant confounders. In many situations, it is desirable to utilize data to infer a treatment's causal effect:

Cherny, Stacey S., [20] et al. has suggested in this document Microbial resistance manifests as cross-resistance and collateral sensitivity, which are patterns of dependence between several drugs. The experimental and clinical environments exhibit different patterns. It's unclear if the variations are due to biological factors or to results that are biased and confounded in therapeutic settings. Our goal was to use clinical data to clarify the underlying dependency patterns between resistance to various antibiotics while taking patient characteristics and prior antibiotic use into consideration. The prevalence of antibiotic resistance in bacterial pathogens is placing an increasing strain on healthcare systems through higher follow-up visits, longer hospital stays, more expensive new medication, and higher expenses associated with mortality and lost productivity. A recent released by the US Centers for Disease Control states that 35,000 people die each year in the USA as a result of antibiotic-resistant diseases, which affect millions of individuals.

M. P. Francino and others [22]. has suggested in this document Antibiotics are overexposed to the human microbiome since they are used in crops and farm animals in addition to medicine. Antibiotic exposure has the ability to quickly change the composition of the microbiome and have an immediate impact on health, such as the selection of opportunistic bacteria that are resistant and can cause acute disease. Long-term effects of antibiotic-induced changes to the microbiome may potentially have an indirect impact on health. The human body's mutualistic microorganisms regulate immunological and metabolic balance as well as a host of other physiological functions. Antibiotic exposure thus has the potential to disrupt numerous fundamental physiological equilibria, hence fostering chronic illness. Furthermore, overuse of antibiotics promotes bacterial resistance, and an overexposed human microbiome has developed into a substantial reservoir of resistance genes, making it harder to control bacterial infections. Reviewing the intricate interactions between antibiotics and the human microbiome, this article focuses on the intestinal microbiota and discusses how using antibiotics affects the microbiota's makeup and function, how antibiotic-induced changes to the microbiota affect immunity, metabolism, and overall health, and how the gut microbiota acts as a reservoir for antibiotic resistance.

In this study, Reed T. Sutton [21] et al. have proposed in today's healthcare, computerized clinical decision support systems, or CDSS, represent a paradigm change. Clinicians can use CDSS to support them in making complicated decisions. Rapid advancements have been made in CDSS since they were initially used in the 1980s. Due to the growing worldwide usage of advanced electronic medical records, they are now frequently administered using computerized clinical processes and electronic medical records. Despite these developments, there are still questions about how CDSS affect patients' results, expenses, and the clinicians that employ them. Over the past ten or so years, many CDSS success stories have been published. However, significant setbacks have also demonstrated that CDSS are not without hazards. In this work, we present a cutting-edge synopsis of clinical decision support system applications in medicine, covering the many kinds, existing applications with demonstrated effectiveness, typical drawbacks, and possible risks. Finally, we offer evidence-based suggestions for reducing risk in the development, application, assessment, and upkeep of CDSS. By augmenting medical choices with specific clinical knowledge, patient data, and other health information, a clinical decision support system (also known as a CDSS) aims to improve the delivery of healthcare. In a traditional CDSS, patient-specific assessments or recommendations are provided to the clinician for consideration after their characteristics are matched to an electronic clinical knowledge base. This process is intended to directly support clinical decision making. These days, clinicians mainly employ CDSSs at the point-of-care, where they integrate their own expertise with the data or recommendations from the CDSS.

3.ANTIBIOTIC EXISTING SYSTEM

As machine learning (ML) algorithms and data gathering techniques progress, so docomputerized decision support systems. They are seldom ever utilized for empirical antibiotic treatment, though. Here, we use machine learning (ML) algorithms on patients' electronic medical records (EMRs) to predict the antibiotic resistance profiles of bacterial infections in hospitalized patients. The data comprised the antibiotic resistance findings of bacterial cultures from hospitalized patients together with their EMRs. To predict antibiotic resistance, we used gradient boosted trees, lasso logistic regression, neural networks, and an ensemble that included all three approaches. Shapely Additive Explanations analysis and permutation testing were used to measure variable influence. We show how machine learning may be used to anticipate antibiotic resistance in bacterial illnesses that hospitalized patients may have. Furthermore, we demonstrate that quick knowledge about the bacterial species causing the infection may significantly enhance forecasts. Physician ought to think about putting such mechanisms in place to support appropriate empirical treatment and maybe lessen antibiotic abuse.

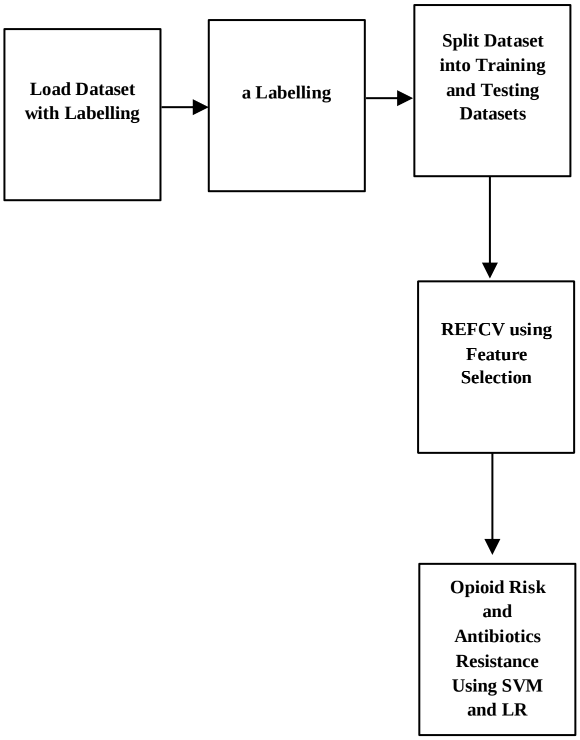

Fig.1 Work flow Diagram

4. ANTIBIOTIC RESISTENCE PREDICTION SYSTEM

The suggested method uses an integrated strategy that makes use of machine learning approaches to solve the complex issues related to opioid risk and antibiotic resistance. Comprehensive datasets containing pertinent characteristics pertaining to opioid usage, associated hazards, and antibiotic resistance factors will be gathered and preprocessed as part of the system. Predictive analytics will use a combination of Support Vector Machine (SVM) and Linear Regression models to improve model interpretability and accuracy. Recursive Feature Elimination with Cross-Validation (REFCV) will be used into the implementation to systematically find and choose the most important features in the datasets. In addition to streamlining the models, this feature selection procedure offers insights into the crucial variables influencing antibiotic resistance and opioid risk. The proposed approach attempts to give healthcare practitioners the ability to forecast the dynamics of antibiotic resistance and opioid-related dangers through rigorous model training and evaluation.

4.1 Load Dataset with Labelling

The suggested system's first step is to load the dataset and give the data the proper labels. This important stage makes sure the data is prepared for further analysis by organizing and structuring it. The dataset is made into a labeled dataset by applying labels, which makes it easier for machine learning algorithms to recognize and extract knowledge from the patterns seen in the data.

4.2 Data Labelling

An essential step in getting the dataset ready for machine learning applications is data labeling. This stage entails giving various data points meaningful labels in order to provide the algorithm context on the target variable. This procedure probably involves classifying cases according to the degree of opioid use and the existence of variables that contribute to antibiotic resistance in the context of managing opioid risk and antibiotic resistance. Accurate and trustworthy prediction models require properly labeled data during training.

4.3 Refcv(Recursive Elimination with Cross-Validation) Using Fearture Selection

Using Recursive Feature Elimination with Cross-Validation (REFCV), the most important features in the datasets are systematically found and chosen. The process of feature selection optimizes the model by eliminating unnecessary variables, cutting down on complexity, and maybe enhancing the interpretability of the model. By verifying the outcomes over several data subsets, cross-validation assures the feature selection's resilience. This stage helps improve the precision and effectiveness of later machine learning models.

4.4 Split Dataset into Training and Testing Datasets

The dataset is divided into training and testing datasets after it has been labeled. This segmentation is essential for evaluating the model's efficacy. The testing dataset, which the model has not seen during training, is used to assess the generalization skills of the machine learning models, which are trained on the training dataset. By ensuring that the model can produce accurate predictions on fresh, untested data, this stage contributes to a realistic evaluation of the model's efficacy.

4.5 Opioid Risk and Antibiotics Resistance Using Svm and Linear Regression

Support Vector Machine (SVM) and Linear Regression models are used in the core predictive analytics phase to address antibiotic resistance and opioid risk. While Linear Regression is useful for modeling interactions between variables and can help understand the mechanisms contributing to antibiotic resistance, Support Vector Machines (SVM) are well-known for their performance in classifying tasks, which makes them excellent for predicting opioid risk categories. By utilizing both models, the suggested approach seeks to give healthcare practitioners a thorough grasp of the dynamics related to opioid use and antibiotic resistance, empowering them to make well-informed decisions.

Support Vector Machine (SVM) is a popular supervised machine learning model that is used for classification and prediction of unknown data. It is asserted by several researchers that SVM is a very accurate technique for text classification. It is also widely used in sentiment classification. For instance, if we have a dataset in which data is pre-labeled into two categories: positive and negative reviews, then we can train a model to classify new data into these two categories. This is exactly how SVM works. It is the model that we train on a dataset, so it can analyze and classify unknown data into the categories that were present in the training set. SVM is a linear learning method. It finds an optimal hyper-plane to differentiate two classes. Being a supervised classification model, it tries to maximize the distance between the closest training point and either class so as to achieve better classification performance on test data. The process for classification functions is as follows:

• It takes the labeled sample of data, and draws a line separating the two classes. This line is called the decision boundary. The solution is based only on those training data points which are really close to the decision boundary. The data points are called Support Vectors. For example, if we are categorizing movie reviews (in our case), one side of the boundary will have positive reviews while the other side has negative reviews.

• Now when new data needs to be classified, it goes either into the left or right side of the decision boundary. Depending on which side the data enters, it is classified under that category. To classify our data with the best precision, we need to split the two categories such that the decision boundary separates the two classes with maximum space between them.

ALGORITHM:

Initialization:

- Initialize the SVM model with parameters such as the regularization parameter(C), kernel type,,kernel parameters(if applicable),and convergence criteria.

Data Preprocessing:

- Standardize or normalize the input features if needed to ensure they are on similar scales.

Select Kernel Function:

- Choose a kernel function based on the problem type and data characteristics.Common choices include linear, polynomial,radial basis function(RBF),or sigmoid kernels.

Define Optimization Objective:

- Formulate the optimization problem that SVM aims to solve.the objective is typically to find the maximum- margin hyperplane that separates thr classes while minimizing classification errors.

Optimization:

- Used method is the Sequential Minimal Optimization (SMO) algorithm,but other optimization techniques such as gradient descent can also be used.

Training:

- Train the SVM model using the training data by optimizing the defined objective function. During training, the algorithm adjusts the model parameters to find the optimal hyperplane that separates the classes with the maximum margin.

Model Evaluation:

- Assess the performance of the trained model using evaluation metrics such as accuracy,precision, recall,F1-score , or area under the ROC curve(AUC)

Parameter Tuning:

- Fine-tune the model parmeters, such as the regularization parameter (C) and kernel parameters,using techniques like cross-validation to improve performance and generalization.

Prediction:

- Use the trained SVM model to make predictions on, unseen data points.The model predicts the class label based on which side of the decision boundary(hyperplane) the data point lies.

Post- processing (if needed):

- Depending on the problem requirements, perform any necessary post-processing steps such as thresholding predicted probabilities or handling class imbalances.

The algorithm outlines the key steps involved in training and using a Support Vector Macine for binary classification.Actual implementations may involve additional optimizations and considerations specific to the problem domain.

B. Linear Regression

Let's get to know about Linear Regression algorithm with the help of an example first, consider a situation wherein, you are required to arrange random logs of work in ascending order of their weight, but there is a condition: you are not allowed to weigh each log, so you have only one option: to guess the weight of each log just by visual analysis. This is exactly the type of problem which can be solved by linear regression. The Linear regression algorithm uses the mathematical equation: $y = bx + a$. This equation describes the line of best fit for the relationship between the dependent variable(y) and the independent variable(x). Following equation governs the Linear regression algorithm

Where Eq. 2 gives the y-intercept(a) and Eq. 3 gives the slope of line(b).

Linear Regression is very significant due to the following reasons:

1. It helps understand the strength of the relationship between the outcome and the predictor variable.

2. Linear Regression can make an adjustment for the effect of covariance.

3. Through Linear Regression, we can get to know the risk factor affecting the dependent variable.

4. It also helps in quantifying new cases in our problem statement.

Regression analysis provides us with three things:

1. Description: Relation between dependent and independent variables.

2. Estimation: Value of the dependent variable can be estimated from observed values on the independent variables.

3. Risk factor: Upcoming risk factor can be predicted beforehand to avoid a critical situation in certain analysis scenarios.

Assumptions for linear regression:

1. Values of the independent variable on the X-axis is set by the researcher.

2. The value of X-axis does not have any experimental error.

3. All the values of Y are independent of each other but depend only on the values on the X-axis that is the dependent variable.

PSEUDOCODE

```
FUNCTION  main():
Step 1: Gather datasets

dataset = gather_datasets()

Step 2: Preprocess the data

preprocessed_data = preprocess(dataset)

Step 3: Feature selection using Recursive Feature Elimination

selected_features
recursive_feature_elimination(preprocessed_data)

Step 4: Train models

svm_model = train_svm_model(selected_features)

linear_regression_model
train_linear_regression_model(selected_features)

Step 5: Combine predictions

combined_predictions combine_predictions(svm_model,
linear_regression_model
```

Step 6: Evaluate model performance

evaluation_results evaluate_model(combined_predictions)

Step 7: Interpret results

interpret_results(selected_features, evaluation_results)

Step 8: Deploy models

deploy_models(svm_model, linear_regression_model)

5. RESULT ANALYSIS

The table shows an accuracy-based overview of algorithmic performance for two different models used in a predictive analytics work. With Lasso logistic regression, the first model achieves a 75% prediction accuracy. Lasso logistic regression is well-known for its feature selection and handling of high-dimensional data. By contrast, the second model achieves a significantly greater accuracy of 88% by combining Linear Regression (LR) and Support Vector Machine (SVM). SVM excels at tackling challenging classification tasks, and LR helps make the model more comprehensible. The table highlights how the combination of SVM and LR performed better in terms of prediction than the Lasso logistic regression model, indicating that the integrated method was effective in producing a more accurate predictive model.

One of the most widely used metrics for assessing classification performance is accuracy, which is calculated as the ratio of correctly segmented samples to all samples.

Accuracy = TP/ (TP+ FN)

Precision: The number of positive class predictions that truly belong to the positive class is quantified by precision, which is estimated in the manner described below.

Precision = TP / (TP + FP)

The ratio of true positives to total (real) positives in the data is known as recall or sensitivity. Sensitivity and recall are synonymous.

Recall = TP / (TP + FN)

The ratio of genuine negatives to total negatives in the data is known as specificity. Specificity is the program's accurate designation for everyone who is actually healthy.

Specificity = TN / (TN + FP)

Algorithm	Accuracy	precision	Recall	F-measure
lasso logistic regression	94.151	87.566	88.318	88.937
SVM and LR	76.115	81.43	78.241	79.804

Table 1. Comparison table

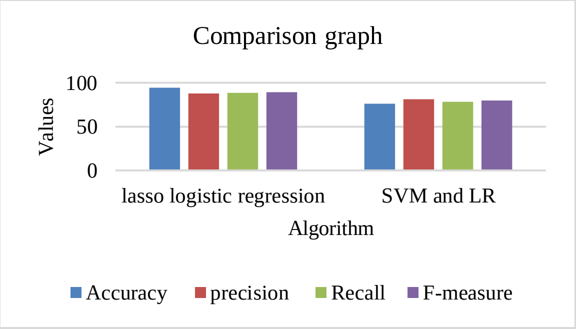

Figure 1. Comparison graph

6. CONCLUSION AND FUTURE WORK

TShe suggested integrated system, which combines Recursive Feature Elimination with Cross-Validation (REFCV), Linear Regression models, and Support Vector Machine (SVM), offers a strong framework for thoroughly tackling the complex issues related to opioid risk and antibiotic resistance. Through the application of machine learning, this system not only makes precise predictions but also reveals the critical elements affecting the phenomena that is being observed. Healthcare practitioners and policymakers may successfully adjust interventions and policies by gaining a deep understanding of the relevant aspects through rigorous feature selection. It is essential that the suggested approach be expanded in subsequent work by adding real-time data sources and updating the models often to guarantee their capacity to adjust to changing patterns in antibiotic resistance and opioid risk. The system's predictive skills might be further improved by integration with new technologies, such as natural language processing for unstructured healthcare data mining and the incorporation of genetic data. Further research into ensemble learning strategies or more complex deep learning architectures could also result in better model performance.

REFERENCES

[1] National Research Council. (1986). *Pesticide resistance: strategies and tactics for management*. National Academies Press.

[2] MacGowan, A., & Macnaughton, E. (2017). Antibiotic resistance. *Medicine*, *45*(10), 622-628.

[3] Anahtar, M. N., Yang, J. H., & Kanjilal, S. (2021). Applications of machine learning to the problem of antimicrobial resistance: an emerging model for translational research. *Journal of clinical microbiology*, *59*(7), 10-1128.

[4] Hernando-Amado, S., Coque, T. M., Baquero, F., Martínez, J. L. (2020). Antibiotic resistance: moving from individual health norms to social norms in one health and global health. *Frontiers in Microbiology*, *11*, 547252.

[5] Chinemerem Nwobodo, D., Ugwu, M. C., Oliseloke Anie, C., Al-Ouqaili, M. T., Chinedu Ikem, J., Victor Chigozie, U., & Saki, M. (2022). Antibiotic resistance: The challenges and some emerging strategies for tackling a global menace. *Journal of clinical laboratory analysis*, *36*(9), e24655.

[6] Pitsavos, C., Panagiotakos, D. B., Lentzas, Y., & Stefanadis, C. (2005). Epidemiology of leisure-time physical activity in socio-demographic, lifestyle and psychological characteristics of men and women in Greece: the ATTICA Study. *BMC public health*, *5*, 1-9.

[7] Turner, P. E. (2013). Evolutionary medicine. *The Princeton Guide to Evolution*, 733-740.

[8] Salam, M. A., Al-Amin, M. Y., Salam, M. T., Pawar, J. S., Akhter, N., Rabaan, A. A., & Alqumber, M. A. (2023, July). Antimicrobial resistance: a growing serious threat for global public health. In *Healthcare* (Vol. 11, No. 13, p. 1946). MDPI.

[9] Baquero, F., Martinez, J. L., F. Lanza, V., Rodríguez-Beltrán, J., Galán, J. C., San Millán, A., ... & Coque, T. M. (2021). Evolutionary pathways and trajectories in antibiotic resistance. *Clinical Microbiology Reviews*, *34*(4), e00050-19.

[10] Livermore, D. M. (2005). Minimising antibiotic resistance. *The Lancet infectious diseases*, *5*(7), 450-459.

[11] Cunha, B. A. (2000). Antibiotic resistance. *Medical Clinics of North America*, *84*(6), 1407-1429.

[12] Jayatilake, S. M. D. A. C., & Ganegoda, G. U. (2021). Involvement of machine learning tools in healthcare decision making. *Journal of healthcare engineering*, *2021*.

[13] Mascaro V, Capano MS, Iona T, Nobile CGA, Ammendolia A, Pavia M. Antibiotic resistance pattern, including methicillin resistance, and prevalence of Staphylococcus aureus carriage among Italian athletes participating in contact sports. Drug Resist Infect 2019; 12:1161–1170.

[14] Diene SM, Baron SA, Rolain J-M. Antibiotic resistance and human microbiomes. 2018; 10:43–52; Human Microbiome J.

[15] Sommer MO, Church GM, Dantas G. A broad pool of antibiotic resistance genes is present in the human microbiome. 2010; Virulence 1:299–303.

[16] Muller-Pebody B, Freeman R, Pouwels KB, et al. Relationship between trimethoprim resistance and the usage of other antibiotics: moving beyond

[17] Costelloe C, Metcalfe C, Lovering A, Mant D, Hay AD. A systematic review and meta-analysis of the impact of primary care antibiotic prescription on patient antimicrobial resistance. 2010; 340:c2096 in BMJ.

[18] Dadgostar P. Antimicrobial resistance: consequences and expenses [6]. 2019; Infect Drug Resist 12:3903.

[19] Wager S, Athey S. Using random forests for the estimation and inference of diverse treatment effects. 2018; J Am Stat Assoc; 113:1228–1242.

[20] Baraz A, Nevo D, Cherny SS, et al. Using Bayesian network modeling, antibiotic cross-resistance trends in

hospitalized patients are revealed. 2020; med Rxiv:dkaa408. At https://doi.org/10.1093/jac/dkaa408, it is accessible.

[21] Francino M. Dysbioses and resistance building in the human gut microbiome in relation to antibiotics. Microbiol Frontiers 2016; 6:1543.

[22]. Sutton RT, Pincock D, Baumgart DC, Sadowski DC, Fedorak RN, Kroeker KI. Clinical decision support systems: an overview: advantages, disadvantages, and successful solutions 2020 NPJ Digit Med; 3:1–10.