Deep learning for the identification of multidrug resistance in MALDI-TOF MS samples of *Escherichia coli*

Xaviera A. López-Cortés
Department of Computer Science and Industries
Universidad Católica del Maule
Talca, Chile
xlopez@ucm.cl

José M. Manríquez-Troncoso
Department of Computer Science and Industries
Universidad Católica del Maule
Talca, Chile
jose.manriquez@alu.ucm.cl

Abstract

Research studying the prediction of antibiotic resistance based on mass spectrometry data and machine learning focuses only on simple models for the identification of resistance to one antibiotic at a time, Even though a problem of multidrug resistance is currently being faced. Therefore, in this study, a multi-label approach for classifying multidrug resistance in *Escherichia coli* samples using raw MALDI-TOF mass spectrometry data and deep learning techniques was developed. The spectra from a recently published public database, encompassing over 4,500 samples of the bacteria under study, were utilized, sufficient for training a deep learning model, specifically a one dimensional convolutional neural network for this case. The use of this architecture proves to be highly efficient, achieving weighted AUROC and AUPRC values equal to or greater than 0.80, as well as a general performance calculated using the Hamming loss metric reaching 0.132. These results demonstrate that the use of deep learning allows for the development of complex models that enable the simultaneous identification of a predefined set of antibiotics, aiding in the determination of a highly effective treatment.

1 Introduction

Antibiotic resistance has emerged as a global health problem in recent years. Only in the United States, over 23,000 people die each year due to this cause [1]. Antibiotic resistance develops when bacteria evolve and can continue to thrive even in the presence of antibiotics that were previously used for their treatment [2]. Two mechanisms of antibiotic resistance can be identified: Intrinsic resistance, which involves bacterial genera or subspecies that possess unique characteristics that provide them with resistance to antibiotics [3]. On the other hand, Acquired resistance occurs when susceptible bacteria develop antibiotic resistance through the adaptation of the genetic code from resistant strains [4], which can occur in three ways:

- Minimization of the intracellular concentrations of an antibiotic.
- Modification of the antibiotic target by genetic mutation.
- Inactivation of the antibiotic by hydrolysis or modification.

37th Conference on Neural Information Processing Systems (NeurIPS 2023).

Currently, the clinical routine to determine if a specific bacterium is resistant to a series of antibiotics requires extracting a sample and subjecting it to a culture in the presence of potential treatments. This test can take between 24 and 72 hours, which is detrimental in severe cases where patients require rapid and accurate treatment. In recent years, the Matrix Assisted Laser Desorption/Ionization Time-Of-Flight Mass Spectrometer (MALDI-TOF) technique has been increasingly incorporated into clinical practice [5] at an exponential rate. This is due to its multiple benefits, such as speed, precision, and minimal requirement of biological material. Therefore, this technique is currently used for microbial identification [6], subspecies identification [7], and analysis of antibiotic resistance [8]. In the area of antimicrobial resistance study, MALDI-TOF is capable of measuring subtle proteomic differences between a susceptible bacterium and one that exhibits resistance to a specific antibiotic. In this regard, research has begun to integrate the information provided by MALDI-TOF mass spectra with machine learning algorithms. Initially, this approach has been used for species or disease identification [9–12], as well as in conjunction with AMR test results to develop models that allow rapid and accurate diagnosis of resistance to a specific antibiotic. Numerous studies have been carried out in this area, mainly focusing on bacteria of major research interest as defined by the WHO [13]. Examples include the use of SVM algorithms to identify vancomycin resistance in Enterococcus faecium samples [14] and to identify methicillin-resistant Staphylococcus aureus derivatives [15]. Among the main issues identified in this field, one of the challenges is defining a standardized methodology for preprocessing spectra, considering the high dimensionality and noise that can be present in this type of data depending on the equipment used for its collection. Researchers have attempted to address this by using kernels to identify peaks with higher variability between classes [16], or by employing autoencoders to reduce mass spectra to a lower-dimensional vector [17]. Additionally, many of the studies only cover a limited number of samples, which hinders the use of more complex deep learning models. However, in 2022, a public database containing over 300,000 MALDI-TOF mass spectra became available [18]. Finally, the majority of the works in this field focus on developing models for the identification of resistance to a single antibiotic, while the current challenge lies in addressing multidrug resistance [19]. The availability of a large number of spectra along with their profiles of antibiotic resistance provides us with the opportunity to develop more complex models that can classify resistance to multiple antibiotics simultaneously. Therefore, in this work, we propose the development of a convolutional neural network model with a multi-label application for identifying multidrug resistance in *Escherichia coli* samples based on raw mass spectra. This approach significantly reduces the typical tasks involved in processing this type of data, such as smoothing, baseline correction, peak picking, among others.

2 Material and Methods

2.1 Dataset

In this study, the mass spectra from the public database DRIAMS [18] have been used, which consists of over 300,000 MALDI-TOF spectra of various bacterial and fungal species, along with more than 750,000 antimicrobial resistance profiles. This database is divided into four subcollections (DRIAMS-A, DRIAMS-B, DRIAMS-C, DRIAMS-D) corresponding to different laboratories where the samples were collected. For this work, the data from DRIAMS-A has been considered, as it has the largest number of samples. For this case, *Escherichia coli* samples has been chosen, which is among the most critical group of priority pathogens identified by the World Health Organization [13], along with the antibiotics Ceftriaxone and Ciprofloxacin. These antibiotics are of high research interest due to the recent increase in identified resistant strains [20, 21]. To construct the datasets used in the models, a matrix was formed using the raw mass spectra, which were subjected to a binning process considering a range of 2,000 Da to 10,000 Da with a bin size of 2 Da to obtain a vector of 4,000 features applicable to our model. Subsequently, the data were normalized between 0 and 1 and finally divided into 70% for training, 10% for validation, and 20% for testing the final model. Table 1 shows the distribution of samples for each of the studied antibiotics.

Table 1: Number of samples for each antibiotic

Antibiotic	***Resistant***	***Susceptible***
Ciprofloxacin	1.449	3.411
Ceftriaxone	1.039	3.821

2.2 Deep learning model

In this work, a one dimensional (1-D) convolutional neural network (CNN) has been implemented for the multi-label classification of resistance or susceptibility to two commonly used antibiotics in the treatment of *Escherichia coli* infections. This type of network is characterized by its ability to learn simple features such as lines or edges when working with images, and to identify more complex features such as subtle differences in peak intensities at a specific mass value that differentiate whether the bacterial strain belongs to one class or another.

A convolutional neural network [22] typically consists of the following elements:

- **Convolutional layers**: A convolution consists of applying a filter (also called kernel) to an input. Repeatedly applying the same filter to different patches of an input results in a map of activations called a feature map, which indicates the locations and strength of a feature detected in input.
- **Pooling layers**: they down-sample their input to reduce their dimensionality in the following layers, save storage space, speed up the computation, and avoid overfitting problems.
- **Dropout**: The term "dropout" refers to dropping out units (hidden and visible) in a neural network. By dropping a unit out, we mean temporarily removing it from the network, along with all its incoming and outgoing connections, commonly used to reduce overfitting.
- **Fully connected layers**: Play the role of classifier in a CNN, is used at the end of the model after the convolutional and pooling layers have performed feature extraction and consolidation.

2.3 Hyperparameter Search

In order to find the model that best fits the problem at hand, the KerasTuner library [23] has been used, which provides a fully configurable hyperparameter optimization framework. For this work, the Bayesian optimization algorithm has been implemented to identify the best combination of the following parameters:

- Number of 1-D convolutional layers (1 to 5).
- Number of filters and kernels for each 1-D convolutional layer (filters: 32 - 128 with a step of 16, kernels: 2 - 11 with a step of 1).
- Number of fully connected layers (1 to 5).
- Number of neurons for each fully connected (32 - 256 with a step of 16).
- Learning rate (0.001 - 0.0001 - 0.00001).

For the 1-D convolution layers and fully connected layers, the RELU activation function was used. In order to mitigate overfitting and expedite the training process, the EarlyStopping callback was implemented, while ReduceLROnPlateau was employed to enhance model convergence. The above implementation was done using tensorflow 2.13, keras 2.13.1 on a nvidia RTX A4000 GPU.

To perform a performance comparison with more traditional machine learning algorithms, a Bayesian hyperparameter search was also performed on the Extreme Gradient Boosting (XGB) algorithm, optimizing the following hyperparameters.

- max_depth (1 to 10).
- min_child_weight (0.0001 to 10).
- max_delta_step (0.0001 to 10).
- gamma (0.0001 to 10).
- eta (0.0001 to 1)

In relation to computational costs, cross validation was not performed in conjunction with the hyperparameter search, so once the best hyperparameters for the neural network and XGB were defined, a 10-fold cross validation was implemented to study the level of generalization of the model. Therefore, the results reported are the mean of 10 iterations (code avaliable: https://github.com/ManriquezJM/Deep-learning-multidrug-resistance).

2.4 Evaluation Metrics

To evaluate the performance of the predictive model, the test set corresponding to 20% of the total data was used. The evaluation metrics for the optimized model were:

AUROC (Area Under the Receiver Operating Characteristic curve): The ROC curve shows the relationship between the true positive rate (sensitivity) and the false positive rate (1-specificity). It is a commonly applied metric for evaluating diagnostic and prognostic models in medicine. In this case, the weighted AUROC of the model and the labels were calculated separately.

AUPRC (Area Under the Precision-Recall Curve): This metric is used to evaluate models where the classes are imbalanced. Unlike AUROC, this metric is calculated considering precision and recall. In this case, the weighted AUPRC of the model and the labels were calculated separately.

Hamming Loss: Commonly used in multi-label classification problems, it is calculated based on the proportion of correctly predicted labels relative to the total number of labels. Hamming Loss ranges from 0 to 1, where a value close to 0 indicates a lower error in label classification.

F1-Score: This metric combines precision and recall to provide an overall measure of model performance for all classes. A F1-Score value near to one indicates better performance.

3 Results

In the present study, a deep learning model was developed for the multi-label classification of antibiotic resistance in *Escherichia coli* bacteria to Ceftriaxone and Ciprofloxacin. The main results obtained are presented below.

3.1 Final deep learning Model

Figure 1 shows the final architecture of the deep learning model that demonstrated the best performance during the hyperparameter tuning process. It consisted of 4 convolutional layers with filter sizes of 64, 80, 32, 80, and kernel sizes of 8, 7, 11, 2, respectively. A Dropout layer was added after the flattening layer to reduce overfitting. Finally, the classification module consisted of 3 fully connected layers with 144, 144, and 192 neurons, respectively.

For the XGB model the best hyperparameters were: max_depth : 8; min_child_weight : 2; max_delta_step : 2; gamma: 0.1 and eta : 0.1

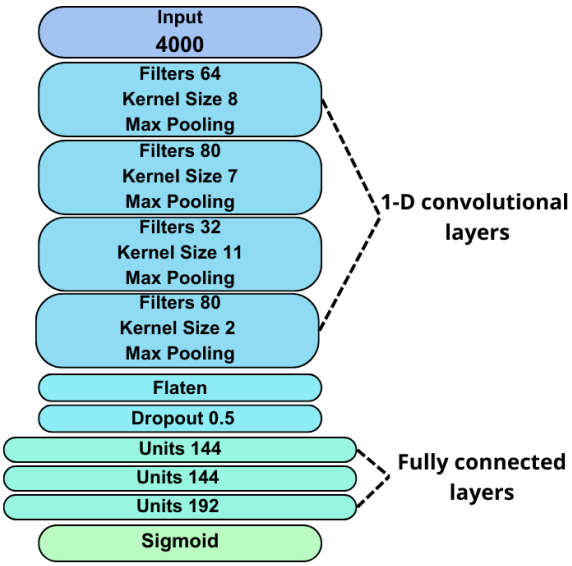

Figure 1: Convolutional neural network model.

3.2 Training performance

As mentioned earlier, various regularization methodologies were applied, such as the use of dropout layers, reducing the learning rate, and early stopping. The training loss and validation loss were plotted for a random split to monitor the model's behavior. Figure 2 shows that with only 45 epochs, the model converges well, achieving a slight difference between the training loss and validation loss of just 0.0496 points.

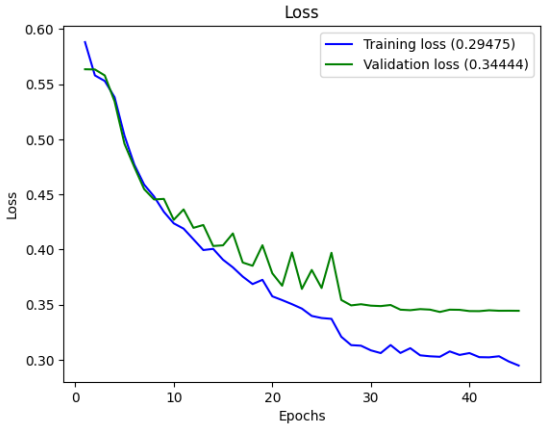

Figure 2: Training history showing the train loss and validation loss for a random split in the deep learning model.

3.3 Classification Performance

To evaluate the performance of the final model, metrics such as AUROC, AUPRC, Hamming Loss, and F1-Score were calculated using the test set. As shown in Table 2, the proposed deep learning model greatly surpassed the performance of the XGB model, achieving a weighted AUROC value of 0.891 and a weighted AUPRC of 0.812, compared to 0.693 (AUROC) and 0.514 (AUPRC) obtained with XGB. Regarding the metrics that assess overall performance across all classes, the Hamming Loss obtained a value of 0.132 in the CNN model, better than the 0.172 achieved by XGB (closer to 0 is better), while the F1-Score reached 0.710, surpassing the 0.535 achieved in the XGB model.

Table 2: 10-fold cross-validation performance results on final models.

Algorithm	***AUROC***	***AUPRC***	***Hamming Loss***	***F1-Score***
CNN	0.891±0.01	0.812±0.03	0.132±0.003	0.710±0.02
XGB	0.693±0.02	0.514±0.03	0.172±0.005	0.535±0.04

The implemented deep learning model also allows us to evaluate each label separately, enabling us to identify which antibiotic performed better individually. As seen in Figure 3, the Ceftriaxone antibiotic achieved a notable AUROC of 0.922 with an AUPRC of 0.849, while the classification of resistance in Ciprofloxacin showed slightly lower performance, with an AUROC of 0.861 and an AUPRC of 0.775.

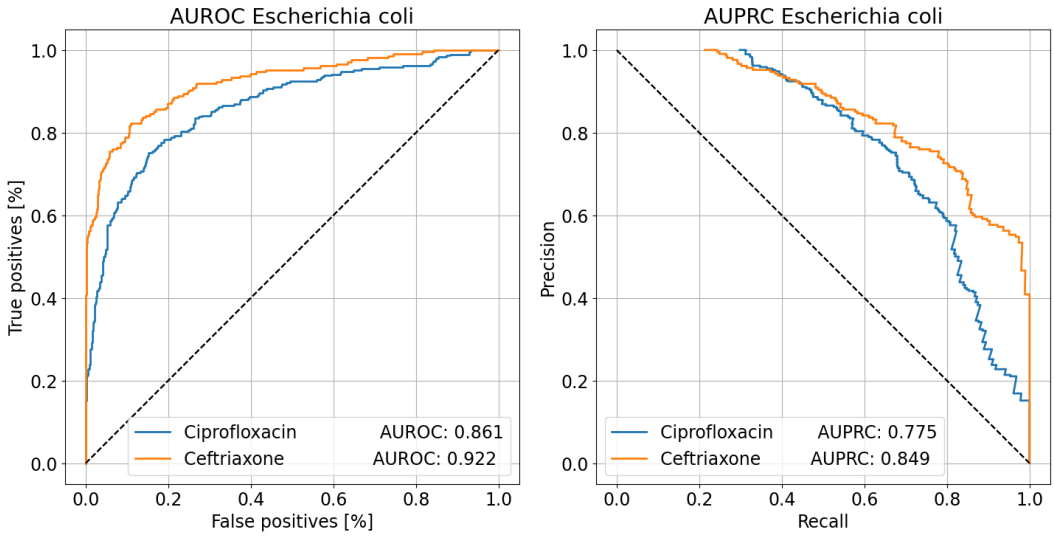

Figure 3: AUROC and AUPRC curves for each of the labels. The value shown in the tables within the figure corresponds to the mean of the 10-fold

3.4 Conclusion and discussion

In this work, a method for the identification of antibiotic multiresistance in *Escherichia coli* bacteria based on the use of convolutional neural networks (CNN) was proposed, from which the following conclusions were drawn.

Regarding spectrum preprocessing, only binning was applied to obtain a fixed-length vector applicable to our models. It was demonstrated that the CNN does not present major issues when working with raw massive data, which can be beneficial in clinical practice as it reduces the time spent on preprocessing tasks. Furthermore, it was shown that the use of deep learning significantly outperforms the results obtained by classical machine learning algorithms like XGB. In terms of overall classification performance, it was observed that training a model capable of discriminating resistance or susceptibility to multiple antibiotics simultaneously is possible. This is crucial for identifying the most effective treatment for an infection. Analyzing the results obtained for each antibiotic, it was found that the resistance classification performed better for Ceftriaxone compared to Ciprofloxacin. This could be directly related to the fact that Ceftriaxone exhibits a greater class imbalance (953 resistant samples vs. 3299 susceptible samples), which is one of the common disadvantages of using neural networks, as they require a large number of samples for training.

In relation to the limitations of this work it is important to mention that the models are adjusted to the spectra collected by a specific MALDI-TOF equipment (Bruker Daltonics) and the bacterial strains were collected in a specific geographic region, so it is important that in future works the compatibility of these models with bacterial strains collected in a totally different geographic region be investigated. On the other hand, to train this type of neural networks it is necessary to have a large amount of data, which are generally not regularly available, either because of the low collection capacity of laboratories or because of privatization policies.

As future work, it is necessary to address the class imbalance in this type of data, and it is also important to investigate how these models can be adapted to mass spectra collected in other geographical areas and with collection equipment from a different manufacturer. This will assist laboratories that have low sample collection capacity but would benefit from the use of a pre-trained model.

In general, it has been demonstrated that the use of CNN is a viable alternative for the identification of antibiotic multiresistance when a large number of samples are available for training. The use of this architecture allows for the construction of complex systems for resistance analysis against a broader range of antibiotics, reducing the time required for treatment decision-making in patient care.

Acknowledgment

The authors gratefully acknowledge financial support from Research Project ANID FONDECYT Iniciación en Investigación No. 11220897.

Funding

This work was supported by ANID Chile, FONDECYT Iniciación grant 11220897.

References

[1] P. Dadgostar, Antimicrobial resistance: Implications and costs, Infection and Drug Resistance 12 (2019) 3903.

[2] M. Skarżyńska, M. Zając, D. Wasyl, Antibiotics and bacteria: Mechanisms of action and resistance strategies, Postępy Mikrobiologii - Advancements of Microbiology 59 (2020) 49–62.

[3] R. E. Impey, D. A. Hawkins, J. M. Sutton, T. P. S. da Costa, Overcoming intrinsic and acquired resistance mechanisms associated with the cell wall of gram-negative bacteria, Antibiotics 9 (2020) 1–19.

[4] A. H. V. Hoek, D. Mevius, B. Guerra, P. Mullany, A. P. Roberts, H. J. Aarts, Acquired antibiotic resistance genes: An overview, Frontiers in Microbiology 2 (2011) 12596.

[5] M. Welker, A. V. Belkum, V. Girard, J. P. Charrier, D. Pincus, An update on the routine application of maldi-tof ms in clinical microbiology, Expert Review of Proteomics 16 (2019) 695–710.

[6] E. Carbonnelle, L. Raskine, Applications de la spectrométrie de masse de type maldi-tof en microbiologie, Bio Tribune Magazine 39 (2011) 35–42.

[7] C. A. Huber, S. J. Reed, D. L. Paterson, Bacterial sub-species typing using matrix-assisted laser desorption/ionization time of flight mass spectrometry: What is promising?, Current Issues in Molecular Biology 2021, Vol. 43, Pages 749-757 43 (2021) 749–757.

[8] M. Cordovana, A. B. Pranada, S. Ambretti, M. Kostrzewa, Maldi-tof bacterial subtyping to detect antibiotic resistance, Clinical Mass Spectrometry 14 Pt A (2019) 3.

[9] C. González, C. A. Astudillo, X. A. López-Cortés, S. Maldonado, Semi-supervised learning for maldi–tof mass spectrometry data classification: an application in the salmon industry, Neural Computing and Applications 35 (2023) 9381–9391.

[10] C. E. Guajardo, X. A. Lopez-Cortes, S. H. Alvarez, Deep learning algorithm applied to bacteria recognition, 2022 IEEE International Conference on Automation/25th Congress of the Chilean Association of Automatic Control: For the Development of Sustainable Agricultural Systems, ICA-ACCA 2022 (2022).

[11] X. A. López-Cortés, F. M. Nachtigall, V. R. Olate, M. Araya, S. Oyanedel, V. Diaz, E. Jakob, M. Ríos-Momberg, L. S. Santos, Fast detection of pathogens in salmon farming industry, Aquaculture 470 (2017) 17–24.

[12] M. T. Ibanez-Barrios, X. A. Lopez-Cortes, X-massfp: A platform with focus on pattern research for mass spectrometry fingerprint recognition, 2021 IEEE International Conference on Automation/24th Congress of the Chilean Association of Automatic Control, ICA-ACCA 2021 (2021).

[13] G. V. Asokan, T. Ramadhan, E. Ahmed, H. Sanad, Who global priority pathogens list: A bibliometric analysis of medline-pubmed for knowledge mobilization to infection prevention and control practices in bahrain, Oman Medical Journal 34 (2019) 184.

[14] H.-Y. Wang, C.-R. Chung, C.-J. Chen, K.-P. Lu, Y.-J. Tseng, T.-H. Chang, M.-H. Wu, W.-T. Huang, T.-W. Lin, T.-P. Liu, T.-Y. Lee, J.-T. Horng, J.-J. Lu, Clinically applicable system for rapidly predicting enterococcus faecium susceptibility to vancomycin, Microbiology Spectrum 9 (2021).

[15] J. Yu, N. Tien, Y.-C. Liu, D.-Y. Cho, J.-W. Chen, Y.-T. Tsai, Y.-C. Huang, H.-J. Chao, C.-J. Chen, Rapid identification of methicillin-resistant staphylococcus aureus using maldi-tof ms and machine learning from over 20,000 clinical isolates, Microbiology Spectrum 10 (2022).

[16] C. Weis, M. Horn, B. Rieck, A. Cuénod, A. Egli, K. Borgwardt, Topological and kernel-based microbial phenotype prediction from maldi-tof mass spectra, Bioinformatics 36 (2020) i30–i38.

[17] Y. Li, Z. Gan, X. Zhou, Z. Chen, Accurate classification of listeria species by maldi-tof mass spectrometry incorporating denoising autoencoder and machine learning, Journal of Microbiological Methods 192 (2022) 106378.

[18] C. Weis, A. Cuénod, B. Rieck, O. Dubuis, S. Graf, C. Lang, M. Oberle, M. Brackmann, K. K. Søgaard, M. Osthoff, K. Borgwardt, A. Egli, Direct antimicrobial resistance prediction from clinical maldi-tof mass spectra using machine learning, Nature Medicine 2022 28:1 28 (2022) 164–174.

[19] R. Giurazza, M. C. Mazza, R. Andini, P. Sansone, M. C. Pace, E. Durante-Mangoni, Emerging treatment options for multi-drug-resistant bacterial infections, Life 2021, Vol. 11, Page 519 11 (2021) 519.

[20] I. Torres, E. Albert, E. Giménez, B. Olea, A. Valdivia, T. Pascual, D. Huntley, D. Sánchez, R. M. Costa, C. Pinto, R. Oltra, J. Colomina, D. Navarro, Performance of a maldi-tof mass spectrometry-based method for rapid detection of third-generation oxymino-cephalosporin-resistant escherichia coli and klebsiella spp. from blood cultures, European Journal of Clinical Microbiology and Infectious Diseases 40 (2021) 1925–1932.

[21] M. D. Phan, K. M. Peters, L. A. Fraga, S. C. Wallis, S. J. Hancock, N. T. K. Nhu, B. M. Forde, M. J. Bauer, D. L. Paterson, S. A. Beatson, J. Lipman, M. A. Schembri, Plasmid-mediated ciprofloxacin resistance imparts a selective advantage on escherichia coli st131, Antimicrobial Agents and Chemotherapy 66 (2022).

[22] Z. Li, F. Liu, W. Yang, S. Peng, J. Zhou, A survey of convolutional neural networks: Analysis, applications, and prospects, IEEE Transactions on Neural Networks and Learning Systems 33 (2022) 6999–7019.

[23] T. O'Malley, E. Bursztein, J. Long, F. Chollet, H. Jin, L. Invernizzi, et al., Kerastuner, `https://github.com/keras-team/keras-tuner`, 2019.