

# Prediction of AMR and antibiotic discovery using Deep Neural Networks: A review

Saanya Yadav, Rashi Sharma, Neha Yadav\*

Department of Biotechnology, Delhi Technological University

Bawana Road, Delhi Technological University, Shahbad Daulatpur Village, Rohini, Delhi, 110042

Department of Biophysics, University of Delhi, Benito Juarez Road, South Campus, South Moti Bagh, Delhi 110021

## Introduction

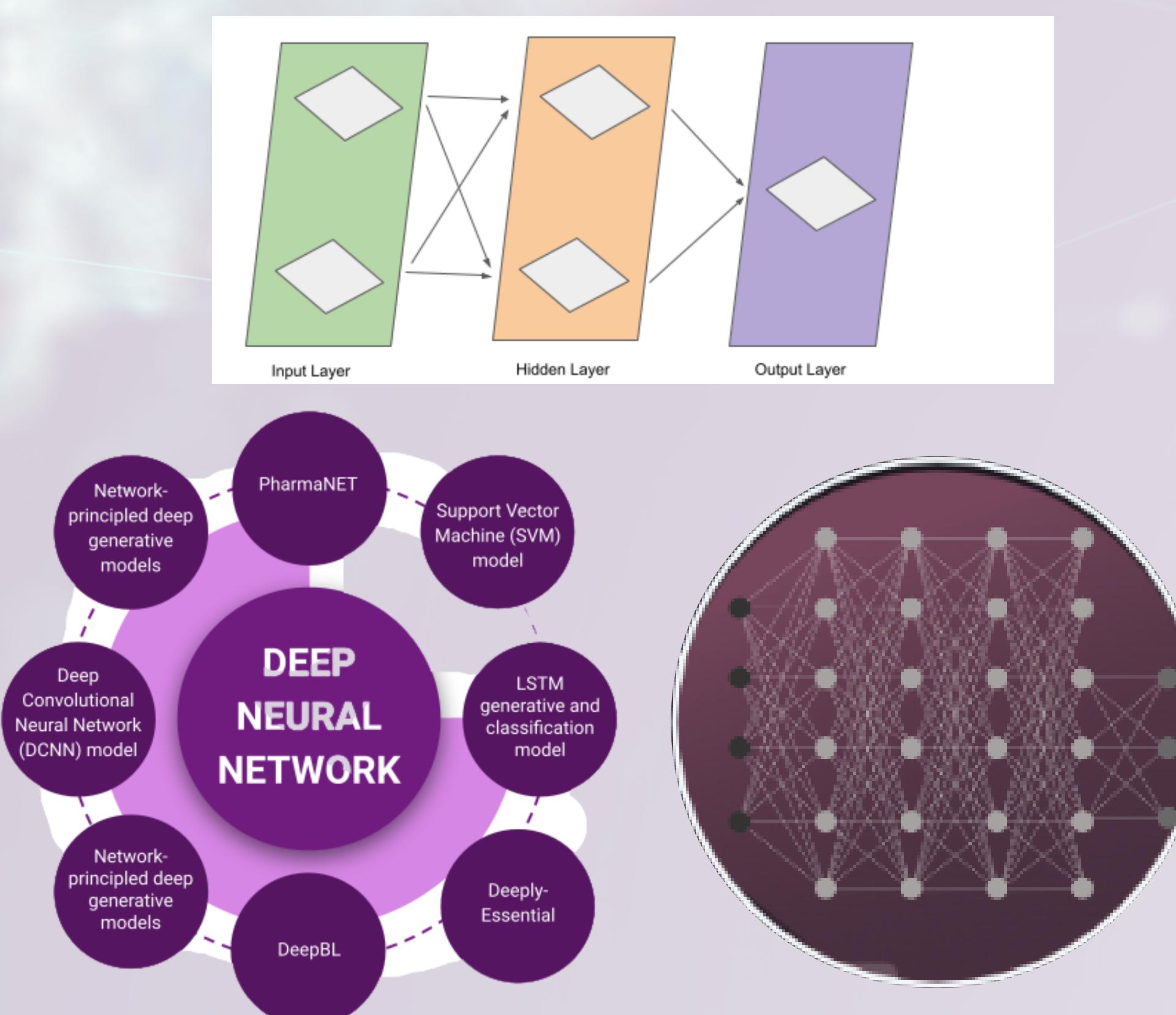
The rise of antimicrobial resistance (AMR) is one of the greatest threats to global health, food security and societal development. Estimates indicate that the number of yearly deaths will be at 10 million worldwide with a cost of \$100 trillion if no steps to tackle AMR are taken by 2050. AMR threatens the effective prevention and treatment of an ever-increasing range of infections caused by bacteria, parasites, viruses and fungi no longer susceptible to the common medicines used to treat them. The problem of AMR is especially urgent regarding antibiotic resistance in bacteria.

A deep neural network can be made capable of predicting molecules with antibacterial activity. Deep learning algorithms also showed significant potential for predicting new antibiotic drugs, AMR genes and AMR peptides. Recently, many studies highlight the potential of machine learning methods in predicting AMR combining sequencing methods and well-known databases with phenotypic information for AMR.

## Working of DNN

Deep Neural Networks (DNNs) are typically Feed-Forward Networks (FFNNs) in which data flows from the input layer to the output layer without going backwards and the links between the layers are one way which is in the forward direction and they never touch a node again.

In order to reveal different sides or features of the input, each layer is capable of performing convolution on the provided data. More the layers deeper the network between input and the output.<sup>3</sup>



## Methodology

The authors of this paper researched various literature archives across the globe like Pubmed, Google Scholar and Scopus for a number of keywords. The selection of reference research articles excluded the studies based on Susceptibility Testing, Pure Machine Learning and Basic Deep Learning. This field of study is relatively new, as the exponential increase in paper is observed from the year 2013. From which the authors thoroughly read and analyzed 45 papers.

## Conclusion

The prevalence of antibiotic resistance is rapidly increasing on a global scale. Concurrently, the steadily declining productivity in clinically implementing new antibiotics due to the high risk of early discovery and low return on investment is INCREASING this problem. Therefore, the development of new approaches that can substantially decrease the cost and increase the rate of antibiotic discovery is essential to reinfuse the pipeline with a steady stream of candidates that show promise as next-generation therapeutics.

Indeed, modern neural molecular representations have the potential to: (1) decrease the cost of lead molecule identification because screening is limited to gathering appropriate training data, (2) increase the true positive rate of identifying structurally novel compounds with the desired bioactivity, and (3) decrease the time and labor required to find these ideal compounds from months or years to weeks.

It is important to emphasize that machine learning is imperfect. Therefore, the success of deep neural network model-guided antibiotic discovery rests heavily on the coupling of these approaches to appropriate experimental designs.

## Existing Models

Model	Architecture	Workflow	Function	Source
PharmaNET	CNN, RNN	Input SMILES → Raw molecular Image → Fingerprint molecular image through CNN → Global Analysis by RNN → Predicts scores of target	Molecule prediction	Ruiz et al., 2021
Word embedding with deep RNN	Skip-gram model, RNN	Input Protein sequences → Through Skip-gram model → Word embedding vectors → RNN	Bacteriocins prediction	Md-Nafiz et al., 2015
Deeply Essential	DNN	Protein coding gene sequence → Feature shortlisting using DNN → Rectified Linear Unit for assigning values in each hidden layer → AUC to check results	Essential Genes prediction	Hasan M et al., 2020
DNP-AAP	DNN	WGS data → Sequence alignment with template → Identify SNP positioning → Feature Selection using DNN → Build Logistic Regression classifiers → Resistance Profiles	Identify known AMR-associated genes in Bacteria	Kusalik et al., 2019
AMP using RNN	RNN	Peptide Database → Peptide Preprocessing → Establish loss function and optimized algorithm → RNN with LSTM cells to generate economical peptide	Synthetic antimicrobial peptides generation	Velez et al., 2021
Neural molecular representations to predict antibacterial compounds	RDKit, SVM, DNN	Screening large chemical libraries (against E. coli BW25113) → molecular graph → ROC-AUC plot → predict scores → Predicted the use of Halicin	Predict molecules with antibacterial activity	Stokes et al., 2021

Model	Architecture	Workflow	Function	Source
LSTM & bidirectional LSTM classifiers	RNN	AMP Data → LSTM and Bidirectional LSTM → Dropout Layer → Formation of Dense and Flatten layer → AMP predictions	Novel antimicrobial peptides formulation	
Gene-centric method for predicting drug resistance in TB	LRCN	PATRIC & ReSeqTB → gene burden features → LSTM+CNN model → Predictions	Predicting Drug Resistance	Safari et al., 2021
Deep learning docking	Deep Learning	Molecular Decoys → Glide SP as DD docking program → Morgan fingerprints → Calculation of Murcko framework	Predicting docking scores of 1.3 billion compounds	Gentile et al., 2020
DeepBL	Deep Learning	Sequence-derived features → Small VGGNet architecture and the TensorFlow deep learning library → model performance evaluation → proteome-wide screening for all reviewed bacterium protein sequences (UniProt database)	Beta-lactamase Prediction	Song et al., 2020
Deep-AmPEP30	CNN and RAAC	PseKRAAC → RAAC Clustering → sequential forward selection approach for model → model tested on genome sequence of <i>Candida glabrata</i>	Improve Short Antimicrobial Peptides Prediction	Siu et al., 2020

## References

- Stokes, J. M., Yang, K., Swanson, K., Jin, W., Cubillos-Ruiz, A., Donghia, N. M., MacNair, C. R., French, S., Carfrae, L. A., Bloom-Ackerman, Z., Tran, V. M., Chiappino-Pepe, A., Badran, A. H., Andrews, I. W., Chory, E. J., Church, G. M., Brown, E. D., Jaakkola, T. S., Barzilay, R., & Collins, J. J. (2020). A Deep Learning Approach to Antibiotic Discovery. *Cell*, 180(4), 688-702.e13. <https://doi.org/10.1016/j.cell.2020.01.021>
- Yunxiao Ren, Trinad Chakraborty, Swapnil Doijad, Linda Falgenhauer, Jane Falgenhauer, Alexander Goesmann, Anne-Christin Hauschild, Oliver Schwengers, Dominik Heider, Prediction of antimicrobial resistance based on whole-genome sequencing and machine learning, *Bioinformatics*, 2021; [btab681](https://doi.org/10.1093/bioinformatics/btab681), <https://doi.org/10.1093/bioinformatics/btab681>
- Wang Y, Li F, Bharathwaj M, Rosas NC, Leier A, Akutsu T, Webb GI, Marquez-Lago TT, Li J, Lithgow T, Song J. DeepBL: a deep learning-based approach for *in silico* discovery of beta-lactamases. *Brief Bioinform*. 2021 Jul 20;22(4):bbaa301. doi: 10.1093/bib/bbaa301. PMID: 33212503; PMCID: PMC8294541.