IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCES
ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

Drug Combination Therapy for Malaria using Deep Learning

Siva Prasad Pinnamaneni[1], Assistant professor, Department of CSE,
Vasireddy Venkatadri Institute of Technology, Nambur, Guntur Dt., Andhra Pradesh.

G. Kamal tej[2], **D. Kartheek**[3], **B. Dayamani**[4], **Ch. Geya**[5], **A. Surendra**[6]
[2,3,4,5] UG Students, Department of CSE,
Vasireddy Venkatadri Institute of Technology, Nambur, Guntur Dt., Andhra Pradesh.
19BQ1A0564@vvit.net, 19BQ1A0542@vvit.net,
19BQ1A0510@vvit.net,19BQ1A0552@vvit.net, 20BQ5A0501@vvit.net

DOI:10.48047/IJFANS/V11/I12/192

Abstract

Drug Combination has been effective for treating complex disorders like cancer and infectious diseases. Malaria remains a major global health challenge, with millions of cases and hundreds of thousands of deaths reported annually. While several drugs are available for malaria treatment, drug resistance has emerged as a significant problem. Combination therapy is now recommended as the first-line treatment for malaria. Due to the impossibility and high cost of considering every possible drug combination, it is necessary to put a lot of work into screening new drug combinations. Recently, deep learning techniques have shown promising results in discovering synergistic combinations. We present synergistic drug combinations for malaria and compare and analyze various models that predict effective drug combinations using deep learning techniques.

Keywords: Malaria, Deep learning, Drug Combination, Drug Synergy.

Introduction

The most recent World Malaria Report shows that in 2021, there were 247 million cases of malaria, which is slightly higher than the 245 million cases reported in 2020. The estimated number of malaria deaths in 2021 was 619,000, compared to 625,000 in 2020. The COVID-19 pandemic caused disruptions that led to an additional 13 million cases of malaria and 63,000 more deaths from the disease over the two peak years of the pandemic (2020-2021). The parasite Plasmodium falciparum, responsible for causing malaria, can develop resistance to various therapies over time through multiple mechanisms. This has resulted in the loss of effectiveness of numerous antimalarial treatments in the past and still poses a significant threat to current standards of care.

The idea behind combination therapy is to utilize the synergistic or additive effects of multiple drugs to enhance the therapeutic efficacy while preventing or delaying the development of resistance to the individual components of the treatment. The use of deep learning[2] has shown enormous potential in the discovery of drug combinations for various

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

diseases. This method involves training neural networks on extensive datasets containing molecular structures and genomic information enabling the prediction of effectiveness of potential drug combinations.

Literature Survey

At the outset for prescribing drugs for simple and trivial diseases Mono-therapy would be sufficient. But for complex diseases like cancer, diabetes, e.t.c. Mono-therapy would not be good and it may lead to some toxic side effects. So, we need a combination of drugs and such complex diseases [9]. For prescribing a combination of drugs we need a lot of manual testing which would be expensive and time consuming. To avoid those many manual trials we can use Machine learning approaches which are used to analyse personalised drug combinations.

In the article by Samuel Kaski and Hiroshi Mamitsuka [7] it is mentioned that the FDA has approved a pair wise drug combination of Dolutegravir and Lamivudine which blocks the HIV-1 multiplication for the treatment of HIV-1 infection.Clinical trials have demonstrated that administering both Vermurafenib, which focuses on BRAF, and Cobimetinib, which targets MAP2K1, leads to a synergistic effect in treating melanoma with BRAF mutation, and this treatment has received FDA approval.

The first thing we need for predicting a combination of drugs for a disease is the chemical properties of each and every drug in the database. Abe Motoki, Mihai Mororiu and Tomoya Otabi have mentioned in their article “Chainer Chemistry: A Library for Deep Learning in Biology and Chemistry” [3] that this can be achieved by using the graph convolution neural network model provided by the Chainer Chemistry. This model takes the molecular structure of a chemical compound and uses the GCNN to predict the biological and chemical properties of the chemical compound.

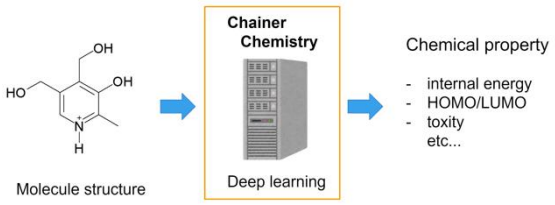

Fig - 1: Illustrating the working of chainer chemistry

Kristina Preuer, Richard I Lewis and Sepp Hochreiter have done a paper on “Deep Synergy: predicting anti-cancer drug synergy with Deep Learning” [1] which shows the usage of DeepSynergy to give more effective results compared to traditional machine learning

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

methods like Support Vector Machines, Elastic Nets and Random Forests on the publicly available synergy dataset.

DeepSynergy was a success in anti-cancer drug prediction. So, we thought of using DeepSynergy for predicting the drug combination for Malaria.

Dr. Yasaman Kalantar and Dr Rajarshi Guha had done an article "Predicting novel combination treatments for malaria using machine learning" [8] which mentions the use of novel combination of drugs to treat malaria and it shows that it 80% of the predicted drug pairs are were synergistic and experimentally validated. This article uses the drugs generated by NCATS and uses machine learning methods to predict novel synergistic compound pairs.

Using this, we thought to use DeepSynergy technique to predict drug combinations which would be more effective on the disease rather than novel combinations. As DeepSynergy uses deep learning which will give more accurate results.

Problem Identification

Malaria is one of the complex diseases caused by parasites transmitted through the bite of infected mosquitoes. The drug treatment to handle the disease is not very effective. Drug combination therapy is a recommended treatment for malaria, as it can increase treatment efficacy and reduce the likelihood of drug resistance. However, identifying the drug combinations can be challenging due to the complex interactions between drugs and the variability of the disease across different regions. [11-19]

Methodology

Drug feature set: For the given set of drugs, Drug feature set contains the molecular feature vector i.e., SMILES refers to the Simplified Molecular Input Line Entry System. Subsequently, SMILES is used for describing chemical species structure using short ASCII strings and edge feature matrices. The data structure consists of nested HashMaps, with the drug feature serving as the key and a corresponding HashMap as the value. The nested HashMaps contain two key-value pairs. One key-pair is a molecular graph created from SMILES that represents the drug and the other key-pair is the feature vector.

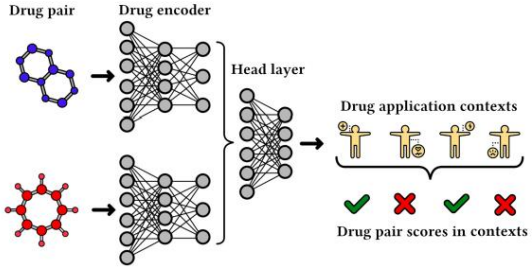

Fig -2: Illustrating the working of the neural network in drug pair combination

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

Context feature set: It refers to a collection of context feature vectors that are used in predicting synergy. These features are utilized to describe genomic expressions.
Drug encoder: It is a type of neural network specifically created to transform the characteristics of a drug molecule into a representation that is low-dimensional and in vector space. This representation can then be used for various drug-related tasks such as drug discovery, virtual screening, and drug-target interaction prediction. Various neural network architectures, including feedforward neural networks and graph neural networks, can be employed by the drug encoder to convert molecular features into a vector space representation.

Context encoder and scoring head layer: The purpose of a context encoder in the field of drug pair scoring is to generate a compressed representation of the biological context, which can then be used as an intermediary input for the drug pair scoring decisions. In deep learning, the scoring head is a crucial component of a neural network as it serves as the final layer responsible for generating a probability or score for the input data. The scoring head layer's output relies on the encoders' representations of both the drug and biological context.

The input data is tested on mainly two models DeepSynergy and MatchMaker. DeepSynergy and MatchMaker outperform well in finding synergistic drug combinations for cancer. DeepSynergy and MatchMaker include the same hyperparameters.

System Implementation

The implementation is mainly done using two techniques[7]. DrugCombdb[6] dataset is used for drug pairs.

DeepSynergy

DeepSynergy[1] is a type of neural network that operates in a feedforward manner, meaning it takes input vectors representing samples and outputs a single value known as the synergy score. These samples are characterized by concatenated vectors that contain the features of two drugs and one cell line.

In the input layer of DeepSynergy, the neurons receive input values which are the gene expression values of the cell line and chemical descriptors of both drugs. These input values are then passed on through the various layers of the DeepSynergy network until they reach the output unit, which produces the predicted synergy score[10].

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

MatchMaker

In the MatchMaker model[5] to predict the synergy score of a drug pair, the model takes in the chemical descriptors of the two drugs and the gene expression profile of an untreated cell line as its features.The model consists of three distinct neural subnetworks. The initial two subnetworks are accountable for acquiring a representation of each drug, which is dependent on gene expression of the particular cell line. The outcomes of these two subnetworks are combined together and passed as input to the third subnetwork, which is responsible for predicting the ultimate output.

Data: $\mathcal{X}_D$ - Drug feature set.
$\mathcal{X}_C$ - Context feature set.
$\mathcal{Y}$ - Labeled drug pair - context triple set.
Result: $\mathcal{L}$ - The cost for the labeled drug pair - context set.

1 $\mathcal{L} \leftarrow 0$
2 **for** $(d, d', c, y^{d,d',c}) \in \mathcal{Y}$ **do**
3 $\quad \mathbf{h}^d \leftarrow f_{\Theta_D}(\mathbf{x}^d, \mathcal{G}^d, \mathbf{X}_N^d, \mathbf{X}_E^d)$
4 $\quad \mathbf{h}^{d'} \leftarrow f_{\Theta_D}(\mathbf{x}^{d'}, \mathcal{G}^{d'}, \mathbf{X}_N^{d'}, \mathbf{X}_E^{d'})$
5 $\quad \mathbf{h}^c \leftarrow f_{\Theta_C}(\mathbf{x}^c)$
6 $\quad \hat{y}^{d,d',c} \leftarrow f_{\Theta_H}(\mathbf{h}^d, \mathbf{h}^{d'}, \mathbf{h}^c)$
7 $\quad \mathcal{L} \leftarrow \mathcal{L} + \ell(y^{d,d',c}, \hat{y}^{d,d',c})$
8 **end**

Deep learning models algorithm to solve drug pair scoring task

□, □′ - Drug pairs

X_N^d and X_E^d - feature matrices

c - Type of context

G^d - Molecular graph

□ □,□′,□ - label drug pair

Line (3),(4) - Drug encoder equation □Θ_D is multivariate parametric function.

Line(5) - Context encoder □Θ_C which is parametrized by $\Theta_{□}$ which depends on the context feature vector $x_{□}$.

Line(6) - The scoring head layer's equation ,denoted as □Θ_H and parametrized by $\Theta_{□}$, is reliant on both drug representations h□ and h□'.

The parameters of the input layer, denoted as Θ_D, consist of biases and weights that are applied to the input data. These parameters are learned through the training process and are utilized to convert the raw input data into a more manageable format for the following layers of the neural network

The parameters of the hidden layers in the neural network are denoted by Θ_C. These layers conduct intermediate computations and facilitate the network in extracting significant features from the input data.

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

The output layer parameters of the neural network are represented by Θ_H. These parameters are utilized to produce the ultimate output of the network, which could be a classification, regression, or other prediction task.

The base Model contains mainly two methods unpack and forward unpack(batch) function returns the context features, left drug features and right drug features. Unpack function is the same for all the models and takes input parameters as a batch. This function unpacks a batch into a tuple of features needed for forward,then the forward() method returns a column vector of predicted synergy scores. DeepSynergy and MatchMaker has same input parameters for forward function as the following syntax ***forward(context_features, drug_features_left, drug_features_right)***

Batch generators are defined to supply batches of drug pairs using labeled drug pair-context triplets, along with sets of drug and context features. A batch of drug pairs is a specialized data structure that includes the drug features, molecular graph, context features, and label for each compound in the batch of drug pairs. The architecture is implemented as PyTorch[4] neural network modules. Models are constructed using PyTorch code without any additional libraries or framework.

Results & Conclusion

In this article, We have tested over 10,000 drug pairs for various cell lines mostly under 3D7 cell_line that derived from the Plasmodium falciparum parasite, which is the most deadly species of the malaria parasite that infects humans and predicted synergy scores.

Some of the predicted synergistic drug combinations are '*Chloroquine* and *artesunate*', '*Artemether* and *lumefantrine*', '*Mefloquine* and *sulfadoxine*'

The performance of the predicted drug combination is measured by the AUC_ROC metric

Model	**Value**
DeepSynergy	0.835±0.002
MatchMaker	0.787±0.002

Future Scope

Deep learning models may not always be able to capture the full complexity of these interactions and lack of flexibility in altering the dosing of individual components. To address this issue, researchers are exploring the use of more sophisticated neural network architectures and integrating domain knowledge into the model.

The future scope is exploring the use of deep learning to identify synergistic drug combinations that have not been previously explored for various other diseases.

References

[1] Kristina Preuer, Richard PI Lewis, Sepp Hochreiter, Andreas Bender, Krishna C Bulusu, and Günter Klambauer. 2018. DeepSynergy: Predicting Anti-Cancer Drug Synergy with Deep Learning. Bioinformatics 34, 9 (2018), 1538–1546.

[2] Benedek Rozemberczki, Stephen Bonner, Andriy Nikolov, Michael Ughetto, Sebastian Nilsson, and Eliseo Papa. 2021. A Unified View of Relational Deep Learning for Drug Pair Scoring. arXiv:2111.02916 [cs.LG]

[3] Abe Motoki, Mihai Mororiu, Tomoya Otabi, Kenshin Abe, and Others. 2017. Chainer Chemistry: A Library for Deep Learning in Biology and Chemistry.

[4] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, et al. 2019. PyTorch: An Imperative Style, High-Performance Deep Learning Library. Advances in Neural Information Processing Systems 32 (2019), 8026–8037.

[5] Kuru Halil Brahim, Oznur Tastan, and Ercument Cicek. 2021. MatchMaker: A Deep Learning Framework for Drug Synergy Prediction. IEEE/ACM Transactions on Computational Biology and Bioinformatics (2021).

[6] Hui Liu, Wenhao Zhang, Bo Zou, Jinxian Wang, and Yuanyuan Deng. 2020. DrugCombDB: A Comprehensive Database of Drug Combinations Toward the Discovery of Combinatorial Therapy. Nucleic acids research 48 (2020), 871–881.

[7] Samuel Kaski, Hiroshi Mamitsuka. Machine learning approaches for drug combination therapies, Briefings in Bioinformatics, Vol. 22, Issue 6, November 2021.

[8] KalantarMotamedi, Y., Eastman, R.T., Guha, R. et al. A systematic and prospectively validated approach for identifying synergistic drug combinations against malaria. Malar J 17, 160 (2018).

[9] Lopez JS, Banerji U. Combine and conquer: challenges for targeted therapy combinations in early phase trials. Nat Rev Clin Oncol 2017.

[10] Pemovska T, Bigenzahn JW, Superti-Furga G. Recent advances in combinatorial drug screening and synergy scoring. Curr Opin Pharmacol 2018.

[11] Sri Hari Nallamala, et al., "A Literature Survey on Data Mining Approach to Effectively Handle Cancer Treatment", (IJET) (UAE), ISSN: 2227 – 524X, Vol. 7, No 2.7, SI 7, Page No: 729 – 732, March 2018.

[12] Sri Hari Nallamala, et.al., "An Appraisal on Recurrent Pattern Analysis Algorithm from the Net Monitor Records", (IJET) (UAE), ISSN: 2227 – 524X, Vol. 7, No 2.7, SI 7, Page No: 542 – 545, March 2018.

[13] Sri Hari Nallamala, et.al, "Qualitative Metrics on Breast Cancer Diagnosis with Neuro Fuzzy Inference Systems", International Journal of Advanced Trends in Computer Science and Engineering, (IJATCSE), ISSN (ONLINE): 2278 – 3091, Vol. 8 No. 2, Page No: 259 – 264, March / April 2019.

[14] Sri Hari Nallamala, et.al, "Breast Cancer Detection using Machine Learning Way", International Journal of Recent Technology and Engineering (IJRTE), ISSN: 2277-3878, Volume-8, Issue-2S3, Page No: 1402 – 1405, July 2019.

[15] Sri Hari Nallamala, et.al, "Pedagogy and Reduction of K-nn Algorithm for Filtering Samples in the Breast Cancer Treatment", International Journal of Scientific and Technology Research, (IJSTR), ISSN: 2277-8616, Vol. 8, Issue 11, Page No: 2168 – 2173, November 2019.

[16] Kolla Bhanu Prakash, Sri Hari Nallamala, et al., "Accurate Hand Gesture Recognition using CNN and RNN Approaches" International Journal of Advanced Trends in Computer Science and Engineering, 9(3), May – June 2020, 3216 – 3222.

[17] Sri Hari Nallamala, et al., "A Review on 'Applications, Early Successes & Challenges of Big Data in Modern Healthcare Management'", Vol.83, May - June 2020 ISSN: 0193-4120 Page No. 11117 – 11121.

[18] Nallamala, S.H., et al., "A Brief Analysis of Collaborative and Content Based Filtering Algorithms used in Recommender Systems", IOP Conference Series: Materials Science and Engineering, 2020, 981(2), 022008.

[19] Nallamala, S.H., Mishra, P., Koneru, S.V., "Breast cancer detection using machine learning approaches", International Journal of Recent Technology and Engineering, 2019, 7(5), pp. 478–481.