International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

A Learning Based Method for the Drug-Drug Interaction Detection

Liya Varghese[1], Lakshmi K.S.[2], Divya James[3], Mathews Abraham[4], Varghese S Chooralil[5]

Submitted: 25/01/2024 **Revised**: 03/03/2024 **Accepted**: 11/03/2024

Abstract: Drug-drug interactions (DDIs) cause grave concern for those patients who require multiple drugs, and in turn for their doctors, caregivers, and society too. Any detection and knowledge imbibed through such interactions utilizing machine learning enables the pharma industry to do away with certain testing modes and helps physicians to impart optimum care while avoiding severe reactions. Here, we put forth a model for predicting any novel drug–drug interaction from a created heterogeneous network, blending in varied drug-relevant information such as drug-disease correlations and drug-side effect correlations, drug–drug interactions etc. which first runs a network diffusion algorithm on each network to determine the "diffusion state," such as random walk with restart. This absorbs its topological relation to other nodes within this diverse network, and forms a drug vector, which is followed by a Denoising Autoencoder model for reducing vector dimensions and identifying vital features. Then, the convolutional neural network model and Support vector classifier is built for predicting drug interactions and evaluating their performances.

Keywords: *Drug-drug interactions, Heterogeneous network, Drug similarity, CNN, Support vector machine*

1. Introduction

The DDIs (Drug-Drug Interactions) are known as the undesirable secondary effects coming about because of the simultaneous utilization of at least two medications. At the point when any specialist endorses a few medications for a certain patient at the same time, DDIs might result in unsalvageable secondary impacts. These may prompt different sicknesses or could even be fatal. These secondary effects are especially recognizable in grown-up individuals and malignant growth patients who consume loads of medications daily. As the significance of anticipating DDIs in human wellbeing, industry and economy, with its measure of cost and conventional trials, the exact computational strategies for foreseeing DDI are out of luck. Vast biomedical data other than the improvement of computational methodologies is available. For instance, DrugBank, a popular and dependable information base of known DDI, contains over 300,000 DDIs. By and by, this measure of communication information is under 1% of the all-out drug matches that exist in DrugBank. Somewhat recently, numerous computational techniques have been created to resolve this issue and beat these impediments. Although past techniques had extraordinary advances, more prediction precision is yet required.

2. Related Work

Currently, numerous computational techniques have been evolved to predict DDIs. DDI forecasting methods can be categorized into text mining, similarity-based prediction and classification-based prediction. Text-mining techniques use natural language processing to extract any potential connections between medications from unstructured data sources. Similarity based strategies presume that comparative medications might interface with a similar medication. For example, two medications might collaborate with the assumption that these have comparable molecular profile. Classification-based procedures simulate the binary classification problem in the DDI prediction task. That is, the presence or lack of collaborations represents drug matching as feature vectors and target variables. Link prediction in classification-based strategies surveys the likelihood of a connection between sets of nodes within a network, considering perception of topology of extant nodes along with traits.

Notwithstanding, scientists recognized a few issues that are disregarded by an extraordinary greater part of DDI expectation studies: (i) powerlessness to foresee recently created drugs, (ii) inability to deal with outrageous information skewness of DDI matches, (iii) dependence on chose information sources (primarily DrugBank), and (iv) reckless assessment procedures that are reflected by utilizing region under ROC bend as the principal assessment metric to evaluate nature of forecast. This multitude of constraints urges to play out a similar report, which assists with finding a superior and further developed experiment for the Medication Cooperation Expectation. A

[1] *Associate QA Automation Engineer, Securiti Sciences Private Limited*
[2] *Associate Professor, Department of Artificial Intelligence and Data Science, Rajagiri School of Engineering & Technology, Kochi, India*
ORCID ID: 0000-0003-1312-3352
[3] *Assistant Professor, Department of Information Technology, Rajagiri School of Engineering & Technology, Kochi, India*
ORCID ID: 0000-0003-3283-9049
[4] *Assistant Professor Department of Information Technology, Rajagiri School of Engineering & Technology, Kochi, India*
ORCID ID: 0009-0002-8368-282X
[5] *Associate Professor, Department of Artificial Intelligence and Data Science, Rajagiri School of Engineering & Technology, Kochi, India*
ORCID ID: 0000-0002-3203-6695
**Corresponding Author Email: lakshmiks@rajagiritech.edu.in*

portion of the current strategies utilized for foreseeing potential DDIs are given beneath.

In the method suggested by Kastrin et.al [1] authors addressed DDIs as one complicated network wherein nodes allude to medications and connections allude to their possible interactions. As of late, 'connection forecast' has drawn in much thought in academic local area. So, they addressed the course of connection forecast as "binary classification" task on any network of expected DDIs. This involved the connection expectation strategies for foreseeing obscure cooperations between drugs in five erratic picked huge scope DDI data sets, specifically Twosides, DrugBank, SemMedDB, KEGG and NDF-RT and assessed presentation of connection expectation by utilizing a set of trials on DDI networks. Next, they executed "link prediction" utilizing "unsupervised" plus "supervised" mode comprising "decision tree, k- nearest neighbors, support vector machine, random forest, and gradient boosting machine classifiers" considering topological and semantic similitude highlights. From the investigations directed, they found that the supervised approach plainly outperformed unsupervised methodology, where area beneath the recall curve for precision was 0.93 for gradient boosting as well as random forest for Twosides network. Thus, they presumed that supervised link prediction way was promising for potential DDIs prediction potential to function with recognizable proof of possible DDIs in clinical examination.

A similarity-based ensemble prediction typel is set up for identifying any potential DDI in the research conducted by Mahadevan et.al [2 where neighbor recommender and Jaccard's coefficient were utilized to calculate similarity measures and random walk algorithm which enhanced prediction through "genetic algorithm" techniques. This ensemble method can identify any drug-drug interaction for varied features. The results showed enhanced precision, recall and accuracy due to random walk and genetic algorithm. Researchers tried this approach to show how this could aid comprehension of DDIs in novel drugs which are clinically administered to any patient who takes other drugs for treating correlative conditions.

Rohani et.al [3] proposed adequate computational procedures to predict unknown DDI with high precision. They used neural network-based procedure for drug affiliation assumption using diverse information about drugs. The "drug substructure, target, side effect, off-label side effect, pathway, transporter, and indication data" were taken into consideration while determining the commonalities between various medicines. All along, they used a heuristic closeness decision cycle, and a while later organized the picked likenesses with "nonlinear similarity fusion" procedure for obtaining irrefutable level features. Thereafter, a neural network was involved for collaboration assumption. Comparability determination and joining bits of NDD have been proposed in past examinations of various issues. So, they combined these parts with a neural network plan and applied the systems concerning DDI assumption.

Takeda et.al suggested a method in which the essential systems of DDIs depended on pharmacokinetics (PK) and pharmacodynamics (PD). The system analyzed impacts of 2D primary similitudes of medications on DDI expectation by collaborating networks of both PD and PK information. Their supposition was that one query drug (Dq) and one medication be inspected (De) possibly having DDI. The assumption was that medications in De interaction network are basically like Dq, where De network portrays relationship between medications and proteins connecting with PK and PD for De. The displaying system contained four stages. In the first phase, association network for every De was built; second, underlying similitudes among Dq and every one of the medications in De network, inc

luding De were figured; third, DDI expectation models were built utilizing primary likenesses with strategic relapse approach; at last, 4-fold cross-validation was conveyed for model assessment.

DDIGIP, a model proposed by Yan et.al for DDI prediction was based on Regularized Least Squares (RLS) classifier and the Gaussian Interaction Profile (GIP) kernel. This kernel is based on drug-drug interaction profiles. The first relational score was calculated using K-nearest neighbors (KNN) and the chemical, biological, and phenotypic information of pharmaceuticals in the presence of novel drugs. Compared to the existing methods (L1 Classifier group strategy), DDGIP strategy obtained AUC values of 0.96 and 0.9636 in 5-fold and 10-fold cross validations respectively, whereas the AUC values were 0.9570 and 0.9599 for the existing methods. Besides, in newer medications, value for AUC of DDIGIP technique arrived at 0.9262 that likewise beats other cutting-edge technique (Weighted normal group strategy) for 0.9073.

3. Proposed Architecture

We hereby recommend a learning-based approach for detecting any drug-drug interaction, which is mainly divided into three modules. The general workflow of this suggested system is shown in Figure 1.

For ease of understanding the proposed model has been divided into 3 modules where the first module is the Heterogeneous-network-based feature extractor, which calculates the similarity matrices for the drug related networks and perform a Random with Restart on these similarity matrices; which is taken as the weighted edge for the network. The Second is the Denoising-Autoencoder (DAE) based feature selector which is used for the dimensionality reduction and the last module is the Drug-

Drug Interaction prediction model, which includes the SVM and CNN model trainings, predictions and their evaluations.

3.1 Network based feature extractor

This heterogeneous network will be built by coordinating an assortment of medication related data sources, comprising drug-drug associations, associations of drugs and their side effects, drug-disease associations and similarities between drugs considering their chemical structure. These datasets are accessible from the publically accessible datasets, for example, DrugBank data set, Similar Toxicogenomic Data set, SIDER data set and so on. First and foremost, the Jaccard similarity calculation [6] is executed on every one of these association and interaction matrix individually and is put away as a text record.

The Jaccard similitude coefficient is a sign of the likeness between two sets which is the comparability between two medications and is defined thus:

$$\boldsymbol{Sim}(\mathbf{A}, \mathbf{B}) = \frac{|\boldsymbol{A} \cap \boldsymbol{B}|}{|\boldsymbol{A} \cup \boldsymbol{B}|} \qquad (1)$$

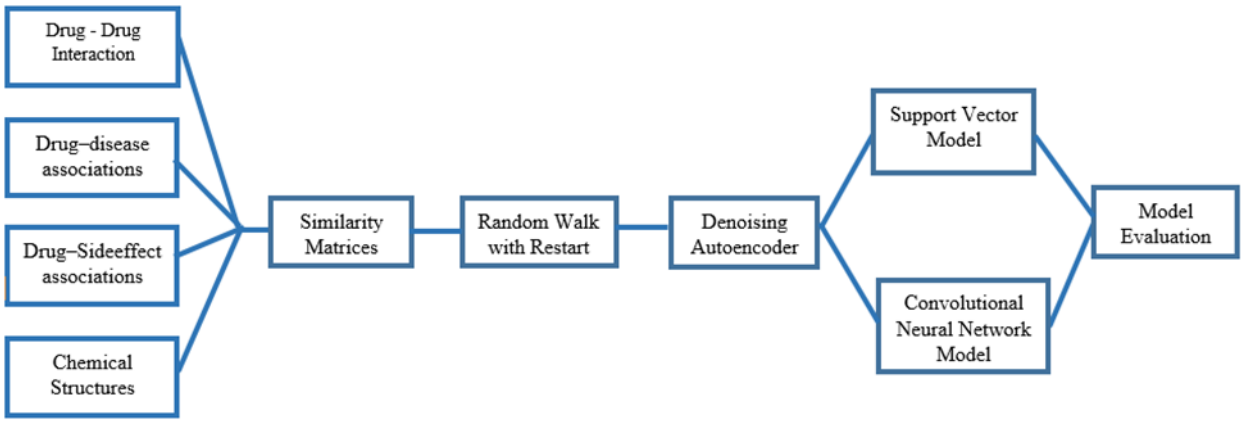


Fig.1: Architecture of the proposed model

Sim_mat_drug_disease.txt

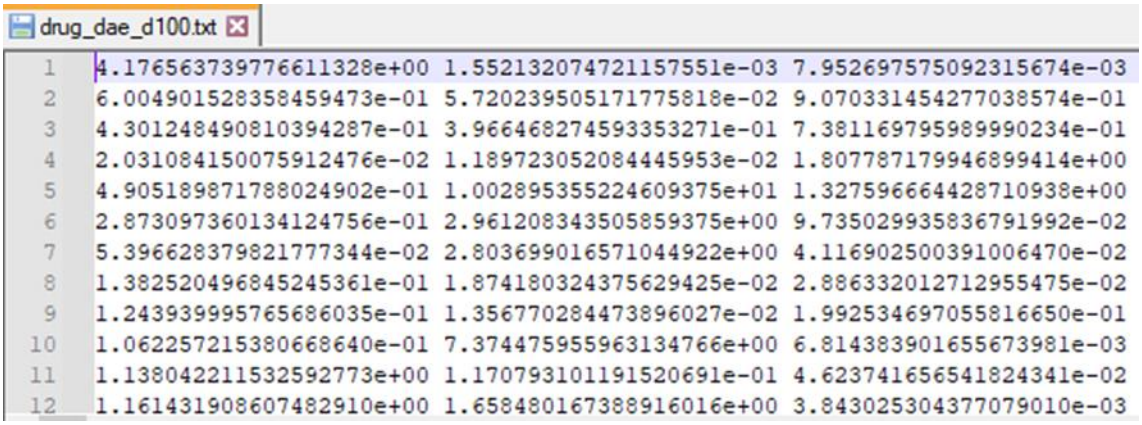
```
1   1   0.19036 0.28817 0.37364 0.18684 0.42368
2   0.19036 1   0.13008 0.20182 0.15571 0.18589
3   0.28817 0.13008 1   0.45246 0.16429 0.45287
4   0.37364 0.20182 0.45246 1   0.22114 0.49258
5   0.18684 0.15571 0.16429 0.22114 1   0.17496
```

Fig.2: Snapshot of drug-disease similarity matrix

drug_vector.txt

```
1   6.5639  0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
2   0   6.5639  0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
3   0   0   6.1515  0.012243    0.017744    0.0079121   0.007606    0.0061863   0   0.0046091
4   0   0   0.043125    5.9478  0.06676 0.036203    0.034032    0.087679    0   0.054914    0
5   0   0   0.2226  0.23803 5.8916  0.45156 0.53973 0.23799 0   0.95707 0   0.40276 0.79801 0
6   0   0   0.079176    0.10156 0.35426 5.9009  2.0439  0.10308 0   0.13244 0   0.066793    0
7   0   0   0.077307    0.097005    0.43236 2.0568  5.9023  0.10429 0   0.10293 0   0.12121 0
8   0   0   0.073929    0.27516 0.21396 0.12172 0.12143 5.896   0   0.1867  0   0.22908 0.322
9   0   0   0   0   0   0   0   0   6.5639  0   0   0   0   0   0   0   0   0   0   0   0   0
10  0   0   0.033716    0.11076 0.61786 0.096523    0.073664    0.1162  0   5.9094  0   0.099
11  0   0   0   0   0   0   0   0   0   0   6.5639  0   0   0   0   0   0   0   0   0   0   0
12  0   0   0.036433    0.083801    0.23043 0.04732 0.085326    0.14107 0   0.097544    0   5
```

Fig.3: Diffusion state matrix of drug features


drug_dae_d100.txt

```
4.176563739776611328e+00 1.552132074721157551e-03 7.952697575092315674e-03
6.004901528358459473e-01 5.720239505171775818e-02 9.070331454277038574e-01
4.301248490810394287e-01 3.966468274593353271e-01 7.381169795989990234e-01
2.031084150075912476e-02 1.897230520844459530e-02 1.807787179946899414e+00
4.905189871788024902e-01 1.002895355224609375e+01 1.327596664428710938e+00
2.873097360134124756e-01 2.961208343505859375e+00 9.735029935836791992e-02
5.396628379821777344e-02 2.803699016571044922e+00 4.116902500391006470e-02
1.382520496845245361e-01 1.874180324375629425e-02 2.886332012712955475e-02
1.243939995765686035e-01 1.356770284473896027e-02 1.992534697055816650e-01
1.062257215380668640e-01 7.374475955963134766e+00 6.814383901655673981e-03
1.138042211532592773e+00 1.170793101191520691e-01 4.623741656541824341e-02
1.161431908607482910e+00 1.658480167388916016e+00 3.843025304377079010e-03
```

Fig.4: Low-dimensional depiction of pharmacological characteristics

When the first information is changed into similarity matrices, it is considered as the weighted edge for the networks and Random walk with Restart [7] calculation is applied to every closeness framework to acquire the diffusion state matrices of each medication in the network and go along with it to shape the single diffusion state network of medication highlights. RWR is used because, in order to take use of any innate direct or indirect linkages among nodes, it introduces a pre-defined restart probability to the first node of each iteration by taking into account "local and global topological connectivity" patterns in a network.

According to this, the greater the nodal similarity, the higher will be its transition probability. [8]. As fewer hyperparameters and lower computation is required for RWR algorithm, this is widely used in complex networks that analyze and feature representation learning [9,10]. RWR algorithm is formulated thus

$$\boldsymbol{x^t = \beta x^{t-1} S + (1-\beta) x^0}, \quad \boldsymbol{\beta} \in (\mathbf{0}, \mathbf{1}) \qquad (2)$$

where x^t is an n-dimensional vector. The j^{th} element of x^t denotes the label confidence score of drug d_j at time step t. x^0 denotes an n-dimensional initial one-hot vector. The j^{th} entry of x^0 will be 1 and all other entries will be 0. M denotes the similarity matrix. M is normalized as $S = T^{-1} * M$, wherein T is diagonal degree matrix with $T_{ii} = \Sigma_j M_{ij}$. β is the restart probability. When L1 norm of $\Delta x = x^t - x^{t-1}$ is less than small positive ε we get a stationary distribution vector x, which is diffusion state of each node [11]. The distribution vectors x will be then stored as matrix X[12].

Subsequently, we splice this for getting a single diffusion state matrix about drug. Columns of drug denotes the 4 nodes corresponding to drug, disease, side effect and any drug and rows denotes varied drugs. The element T(i,j) stands for transition probability between node j and drug i.

3.2 Dimensionality Reduction

Vector of diffusion state matrix generated in the former module is high-layered, and therefore inadequate. For obtaining fundamental elements, we apply a Denoising-Autoencoder (DAE) model [18] that executes information procedure based on autoencoder. Based on input data, autoencoders use programmed encoders to obtain low-dimensional information via "neural networks." Essentially, the Decoders recuperate the unique contribution from low-dimensional information[13]. So, in Denoising-Autoencoder-based feature selector model, to get the low-dimensional representations of those drugs, a noise factor is added to the input vector and the model will learn the low dimensional features from code produced by encoder in the Autoencoder and can obtain the vital aspects from original input for a more robust representation. In this method, DAE has one concealed layer having 100 units and each batch consists of 16 samples. 20 is taken as the number of epochs. Thus, drug feature dimension is reduced to 100 from the original dimensions, which is 2832 and noise figure is set as 0.2, [18] utilizing softplus [14] and RMSProp function [15] to optimize mean-square error (MSE) [16]. Finally, backpropagation (BP) algorithm helps guide the DAE [17][18] which provides the low dimensional vector representation of given input.

Finally, the model embraced 10-fold cross validation technique for isolating train and test set. Here, 90% of known and unknown examples have been utilized for model preparation and 10% examples were acquired for model testing. As indicated by the known drug-drug interaction matrix, we randomly chose the unknown examples with same number of positive or known examples. Altogether, we have 20,000 examples. Subsequent to joining the drug vectors, drug pair vectors of 200 aspects are obtained, which is then given to the for building a prediction model and their exhibitions are assessed.

3.3 Drug-Drug Interaction Prediction

We used CNN as a supervised learning model for predicting drug interaction, influenced by CNN's success in classification tasks[19]. This prediction model contains four layers; the "convolution layer" - which helps model imbibe any local or global structure from input vector[20]. Here the convolutional layer was comprised of 4 kernels with "rectified linear unit (ReLU)" activation function [19] as feature extractor [20], to efficiently ease up calculations, by avoiding gradient explosion and disappearance. After this, the "max-pooling layer", decreased feature map dimension [21], while the pooled size was taken as $2*1$, and the step size was 2 in our model. Then, a one-dimensional vector connected vital features extracted from all kernels. These were passed to the "fully connected layer" with 180 hidden units and an output unit of sigmoid layer which was created for "binary classification" of drug interaction prediction [18].

To compare the models and identify the optimum model among them, we have also used a support vector machine for predicting these drug interactions. SVM is a supervised machine learning classifier, which can map input data set into high-dimensional feature space and later goes on to construct a hyperplane to segregate classes according to maximum margin principle. Several kernel functions exist like linear/nonlinear kernels [25]. Here we have employed the Radial Basis Function (RBF) as the kernel for drug-drug interaction prediction, which is the most generalized one due to its similarity to the Gaussian distribution.

4. Results & Discussions

For implementing the model, the heterogeneous network was built by collecting three kinds of nodes, the drug, disease and the side-effects and three types of edges, the 'drug-drug interaction', 'drug-disease association' and 'drug-side-effect association'. 'Drug-drug interactions' taken from DrugBank database [22] constitutes the drug nodes while 'drug–disease associations' taken from 'Comparative Toxicogenomic Database' [23] forms the disease nodes. The "SIDER database" was used to compile the side-effect nodes and drug–side-effect connections. [24].

First, we constructed four similarity matrices for drug relevant networks based on Jaccard's similarity coefficient. Fig.2 shows a snapshot of drug similarity matrix.

Then a Random walk with Restart algorithm is performed on these similarity matrices, where similarity was taken as the weighted edge for the network. The model performed RWR algorithm for drug related similarity matrices respectively, spliced into a single diffusion state matrix of drug network. Fig.3 gives a snapshot of diffusion state matrix of drug features. Rows of drug diffusion matrix denote diverse drugs, while columns denote proteins, diseases, side effects and drug nodes. The matrix values denote any association between drugs and 4 biological entities. This vector is noisy, high-dimensional and incomplete. For obtaining essential features, a DAE model was applied to carry data operation, based on the Autoencoder. Snapshot of result of the DAE model application is depicted in Fig.4. where in the dimension of drug features is reduced to 100.

Finally, the Drug-Drug Interaction prediction model was constructed using SVM and CNN models. In CNN model, the dropout layer is added before fully connected layer. Dropout percentage was 0.5. The Adam algorithm was used and for the optimization of binary cross entropy loss, the initial learning rate was set as 0.001.

Here, we adopted 10-fold cross validation in which 90% of samples including positive as well as negative samples were considered for training model and 10% for testing. The known drug-drug interactions pairs were taken as positive samples and randomly selected negative samples whose count is equivalent to the count of positive samples. Altogether there were 20,000 samples. After joining the drug vectors, drug-drug pair vectors of 200 dimensions are obtained which was then given to CNN and SVM model for building a prediction model and their performances are thereafter evaluated and compared.

```
new_model.summary()
```

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
convolution_1d_layer (Conv1D (None, 200, 4)            20
_________________________________________________________________
max_pooling_layer (MaxPoolin (None, 100, 4)            0
_________________________________________________________________
reshape_layer (Flatten)      (None, 400)               0
_________________________________________________________________
dropout_layer (Dropout)      (None, 400)               0
_________________________________________________________________
full_connect_layer (Dense)   (None, 128)               51328
_________________________________________________________________
dense (Dense)                (None, 1)                 129
=================================================================
Total params: 51,477
Trainable params: 51,477
Non-trainable params: 0
_________________________________________________________________
```

Fig.5: CNN Model Summary

```
Confusion Matrix
tf.Tensor(
[[9991    9]
 [  14 9986]], shape=(2, 2), dtype=int32)
Accuracy for Predicted-Drug-Interaction : 99.885
Precision for Predicted-Drug-Interaction : 99.90995497748875
Recall for Predicted-Drug-Interaction : 99.86
F1 score for Predicted-Drug-Interaction : 0.998849712428107
```

Fig.6: Confusion matrix of the CNN model prediction with 20,000 samples

```
Confusion Matrx
[[9999    1]
 [  91 9909]]
Accuracy: 0.9954
Precision: 0.999899091826438
Recall: 0.9909
Average_precision: 0.9953500100908175
Aupr score: 0.997674545913219
Auc score: 0.9954
```

Fig.7: Confusion matrix of the SVM model prediction with 20,000 samples

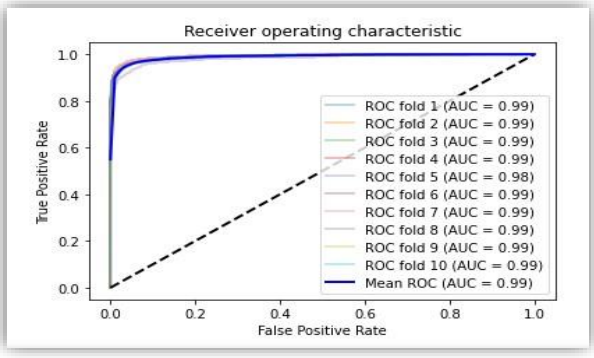


Fig.8: AUROC score for the CNN Model with 10-fold cross validation

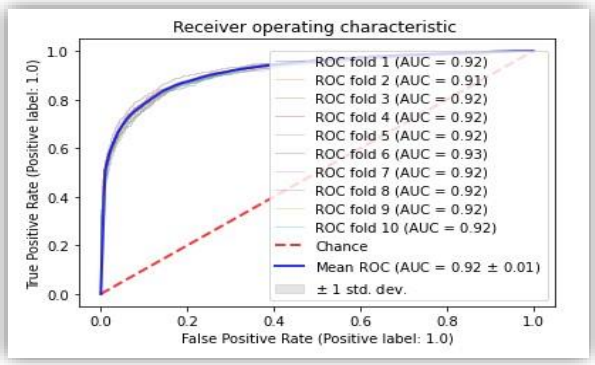


Fig.9: AUROC score for the SVM model with 10-fold cross validation

Performance evaluation

After we built our model, "AUROC" and "AUPR" scores helped in evaluating the model that we had trained and tested. We have used the 10- fold cross validation for model evaluation on both the CNN and SVM models. From the 20,000 samples, we split the samples to 18,000 training data and 2000 testing data and performed the cross validation. Then we found the mean AUROC and AUPR scores for which prediction models have been built. They denote areas under ROC curve and PR curve respectively. These scores are used since they are the common evaluation criteria for machine learning. The higher their score value, the higher will be the accuracy of prediction and hence the prediction data performance [18]. Figure 5 depicts the summary of CNN model. Confusion matrices of the prediction with 20,000 samples using CNN model and SVM model are shown in figures 6 and 7 respectively. Figure 8 shows the evaluations scores when the CNN model is tested on the dataset that is used for training the model. The performance of DDI prediction with the SVM model is shown in figure 9.

From the above result, we can see that the CNN model has better scores when evaluated. Then the CNN model has been tested to predict the drug-drug interaction on entirely new data which is having drug interactions and non-interactions. The model was able to correctly classify that drug combination as interacting and non-interacting based on the input data. The figure 4 shows the AUROC for the CNN Model with 10cross validation and Figure 5 shows the AUROC for the SVM model with 10-cross validation. Since the CNN has better performance, it is chosen as the best model for Drug-Drug Interaction Prediction.

5. Conclusion and Future Scope

In this research, we suggested a learning-based method for drug-drug interaction predictions. Here, Jaccard closeness coefficient and RWR, are the first and foremost model that are utilized to acquire the important elements

of medications from heterogeneous network. Then, at that point, to reduce dimensions and to distinguish the fundamental highlights, we utilized a DAE mode. Thirdly, in view of the highlights got from the past module, the CNN and SVM model were developed. Also, in light of their exhibition assessment we arrived at the resolution that the CNN has better interaction prediction capability when contrasted with the SVM model.

Henceforth, we consider adding up more significant data to this heterogeneous network. We can add network setup for this CNN model that is suitable for accommodating intricate input networks. Here, we have predicted DDIs, as it is an extendible strategy and can likewise be utilized to anticipate other related bearings later on, for example, drug-target interactions, drug-side effects etc.

References

[1] Kastrin A, Ferk P, Leskoˇsek B. Predicting potential drug-drug interactions on topological and semantic similarity features using statistical learning. PLOS ONE 13(5):e0196865, (2018).

[2] Mahadevan, A. A., Vishnuvajjala, A., Dosi, N., Rao, S.: A Predictive Model for Drug-Drug Interaction Using a Similarity Measure. IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), pp. 1-8. (2019)

[3] Rohani, N., Eslahchi, C. Drug-Drug Interaction Predicting by Neural Network Using Integrated Similarity. Sci Rep 9, 13645 (2019).

[4] Takeda T, Hao M, Cheng T, Bryant SH, Wang Y. Predicting drug-drug interactions through drug structural similarities and interaction networks incorporating pharmacokinetics and pharmacodynamics knowledge. J Cheminform. (2017).

[5] Yan, C., Duan, G., Pan, Y. et al. DDIGIP: predicting drug-drug interactions based on Gaussian interaction profile kernels. BMC Bioinformatics 20, 538 (2019)

[6] Niwattanakul S, Singthongchai J, Naenudorn E, Wanapu S. Using of jaccard coefficient for keywords similarity. In: Proceedings of the International Multiconference of Engineers and Computer Scientists; 2013. p. 380–4.

[7] Tong H, Faloutsos C, Pan J-Y. Random walk with restart: fast solutions and applications. Knowl Inf Syst. 2008;14(3): 327–46

[8] Lee S, Lee J, Lim J, Suh I. Robust stereo matching using adaptive random walk with restart algorithm. Image Vis Comput.2015;37:1–11. https://doi.org/10.1016/j.imavis.2015.01.003.

[9] Yan, X. Y., Zhang, S. W., and Zhang, S. Y. (2016). Prediction of Drug-Target Interaction by Label Propagation with Mutual Interaction Information Derived from Heterogeneous Network. Mol. Biosyst. 12 (2), 520–531. doi:10.1039/c5mb00615e

[10] Luo, Y., Zhao, X., Zhou, J., Yang, J., Zhang, Y., Kuang, W., et al. (2017). A Network Integration Approach for Drug-Target Interaction Prediction and Computational Drug Repositioning from Heterogeneous Information. Nat. Commun. 8 (1), 573–613. doi:10.1038/s41467-017-00680-8

[11] Cho, H., Berger, B., and Peng, J. (2015). Diffusion Component Analysis: Unraveling Functional Topology in Biological Networks. Res. Comput. Mol.

[12] Biol. 9029, 62–64. doi:10.1007/978-3-319-16706-0_9

[13] Yan, X. Y., Yin, P. W., Wu, X. M., & Han, J. X. (2021). Prediction of the Drug-Drug Interaction Types with the Unified Embedding Features from Drug Similarity Networks. Frontiers inpharmacology,12,794205. https://doi.org/10.3389/fphar.2021.794205

[14] Peng J, Guan J, Shang X. Predicting parkinson's disease genes based on node2vec and autoencoder. Front Genet. 2019; 10:226. https://doi.org/10.3389/fgene.2019.00226

[15] Ramachandran P, Zoph B, Le Q, Quoc V. Searching for activation functions. arXiv e-prints. 2017. https://ui.adsabs.harvard.edu/abs/2017arXiv171005941R. Provided by the SAO/NASA Astrophysics Data System

[16] Mukkamala M, Hein M. Variants of RMSProp and Adagrad with logarithmic regret bounds. In: Doina P, Yee Whye T, editors. Proceedings of the 34th International Conference on Machine Learning. vol. 70. Sydney: PMLR; 2017. p.2545–53

[17] Allen D. Mean square error of prediction as a criterion for selecting variables. Technometrics. 1971;13(3):469–75.

[18] LeCun Y, Boser B, Denker J, Henderson D, Howard R, Hubbard W, Jackel L. Backpropagation applied to handwritten zip code recognition. Neural Comput. 1989;1(4):541–51

[19] Peng, J., Li, J. & Shang, X. A learning-based method for drug-target interaction prediction based on feature representation learning and deep neural network. BMC Bioinformatics 21, 394 (2020). https://doi.org/10.1186/s12859-020-03677-1

[20] Clevert D-A, Unterthiner T, Hochreiter S. Fast and accurate deep network learning by exponential linear units (elus). arXiv e-prints. 2015arXiv:1511 ˙ 07289.

https://ui.adsabs.harvard.edu/abs/2015arXiv151107289C

[21] [20] Luo P, Ding Y, Lei X, Wu F. deepdriver: predicting cancer driver genes by convolutional neural networks. Front Genet. 2019;10:13

[22] Peng J, Hui W, Li Q, Chen B, Hao J, Jiang Q, Shang X, Wei Z. A learning-based framework for miRNA-disease association identification using neural networks. Bioinformatics.2019;35(21):4364–71. https://academic.oup.com/bioinformatics/article-pdf/35/21/4364/30330838/btz254.pdf.

[23] Knox C, Law V, Jewison T, Liu P, Ly S, Frolkis A, Pon A, Banco K, Mak C, Neveu V, et al. Drugbank 3.0: a comprehensive resource for omics research on drugs. Nucleic Acids Res. 2010;39((suppl_1)):1035–41.

[24] Davis A, Murphy C, Johnson R, Lay J, Lennon-Hopkins K, Saraceni-Richards C, Sciaky D, King B, Rosenstein M, Wiegers T, et al. The comparative toxicogenomics database: update 2013. Nucleic Acids Res. 2012;41(D1):1104–14.

[25] Kuhn M, Campillos M, Letunic I, Jensen L, Bork P. A side effect resource to capture phenotypic effects of drugs. MolSyst Biol. 2010;6(1):343.

[26] Friedman J, Hastie T, Tibshirani R. The elementsof statistical learning:Data mining, inference, and prediction. New York,NY:Springer;2001.