# Drug-Drug interaction predicting by neural network Using integrated Similarity

Narjes Rohani, Changiz Eslanchi

# Novelty:

1. Multiple drug similarities matrices based on different characteristics of the drugs are taken in account.

2. Select the best subset of similarity matrices.

3. Applying SNF, a fusion method, to integrate all selected similarity matrices into an m*m matrix where m is the number of drugs.

4. For each pair of the drugs, their feature vectors are concatenated in a vector and this is considered as the input of a neural network for classification.


# Similarity Selection:

# Similarity selection:

A. **Calculating drug similarities for each drug pair and GIP on interaction matrix only for training data:**

Let matrix $Y = [y_{ij}]$ be the interaction matrix that $y_{ij} \in \{0, 1\}$

$$GIP_d(d_i, d_j) = exp(-\gamma_d ||Y_i - Y_j||^2)$$

B. **Calculating entropy of each matrix:**

$A = [a_{ij}]_{m \times m}$ similarity matrix between $m$ drugs. The entropy $E_i(A)$ for $i$th row is

$$E_i(A) = -\sum_j p_{ij} \log p_{ij} \qquad \text{where} \qquad p_{ij} = \frac{a_{ij}}{\sum_k a_{ik}}$$

# Similarity selection:

C. **Rank matrices based on their entropy in ascending order.**

Matrices with values greater than $c_1\log(m)$ are removed.

D. **Calculate the pairwise distance of matrices.**

The similarity between two feature matrices $A$ and $B$ is

$$S(A, B) = \frac{1}{1 + D(A, B)} \qquad \text{where} \qquad D(A, B) = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{m}(a_{ij} - b_{ij})^2}$$

# Similarity selection:

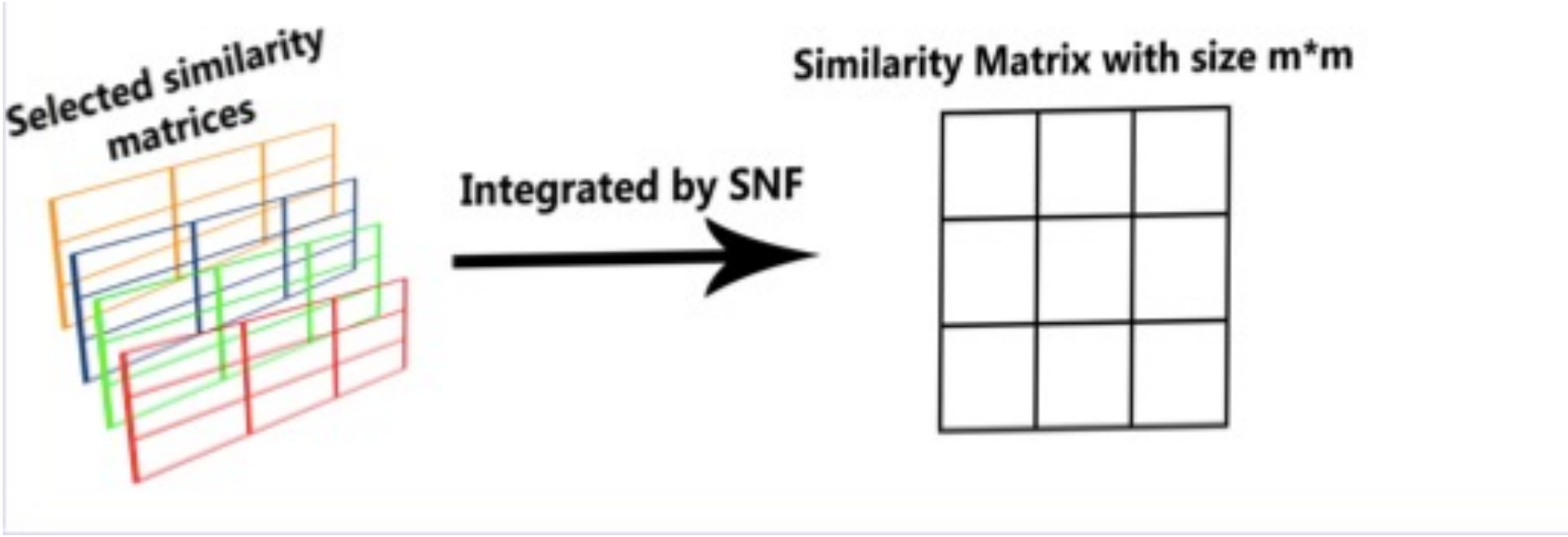
## E. Final selection based on redundancy minimization.

Iterative procedure:

Until the list of ranked matrices based on entropy is empty:

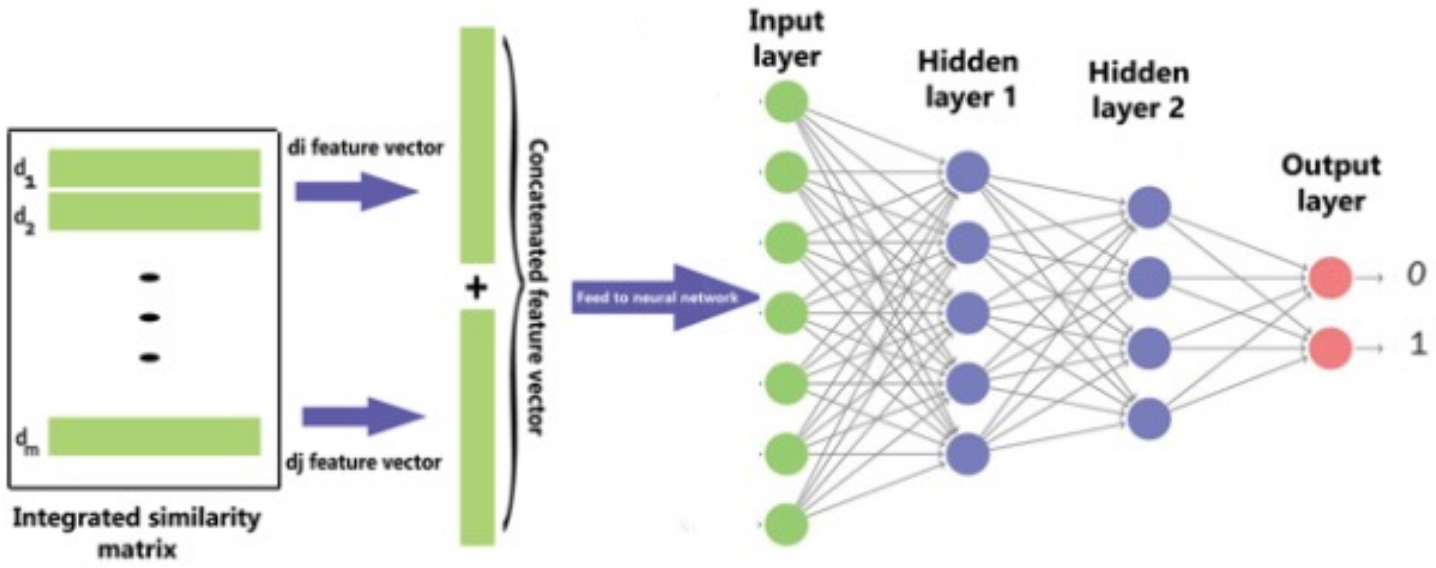
Select the matrix with the least value of the entropy (highly informative) $C=\text{argmin}_i E(A_i)$

Eliminate all matrices $A_j$ , from the list, that have great similarity with $C$,$S(C, A_j) > c_2$ (low redundancy)




**Similarity Network Fusion:**

SNF applies an iterative nonlinear method that updates every similarity matrix according to the other matrices via KNN.




## Neural Network for classification:

Each possible pair of drugs are concatenated and considered as a feature vector in the input layer of the neural network.

# Neural network for classification:

**Nested cross validation to tune the hyperparameters:**

1. Setting the hyperparameters to some values.
2. Partitioning the data into three "folds" (sets).
3. Training the model using two folds with the hyperparameter values.
4. Testing the model on the remaining fold.
5. Performing steps 3 and 4 again and again; thus, every time one of the folds is considered as the test data.
6. Repeating steps 1 to 5 for all combinations of hyperparameter values.
7. Returning the combination of hyperparameter values that had the best performance.

# Neural network for classification:

**Hyperparameter sets:**

- Number of hidden layers: {1, 2, ..., 5}
- Number of neurons in hidden layers: {100, 200, 300, 400, 500}
- Activation functions: {Rectified linear activation function (ReLU), hyperbolic tangent (tanh), and sigmoid}
- Dropout rate: {0.3, 0.5}

# Neural network for classification:

**Train the model:**

- Batch size and epoch number.
- Initialization of the weights and biases.
- Surrogate loss function.
- Optimization algorithm.
- Stratified K-fold cross validation.

# Result:

Datasets Designed from:

- *DrugBank*
- *SIDER*
- *KEGG*
- *PubChem*
- *OFFSIDES*

| Dataset Name | DS1 | DS2 | DS3: CYP | DS3: NCYP |
|---|---|---|---|---|
| N° Drugs | 548 | 707 | 807 | 807 |
| N° of Pairs | 300304 | 499849 | 651249 | 651249 |
| N° of Interactions | 97168 | 34412 | 10078 | 40904 |
| N° of Non-interactions | 203136 | 465437 | 641171 | 610345 |
| N ° of Similarities | 8 | 1 | 7 | 7 |
| Similarity Types | Chemical, Target, Transporter, Enzyme, Pathway, Indication, Side effects, Offside effect | Chemical | GO, Target, Ligand, Chemical, PPI Distance, Side effect, ATC | GO, Target, Ligand, Chemical, PPI Distance, Side effect, ATC |
| AUPR | 0.922 | 0.890 | 0.830 | 0.947 |
| F-measure | 0.835 | 0.825 | 0.772 | 0.902 |
| Recall | 0.836 | 0.804 | 0.770 | 0.884 |
| Precision | 0.833 | 0.847 | 0.775 | 0.918 |

# Strengths & Weaknesses of the procedure:

+ It takes advantage of diverse information from different aspects of drug.
+ It utilizes a similarity selection approach and a fusion procedure to avoid noise, reduce redundant information and exclude random-like data.
+ Neural Network performance was significant in comparison with the graph-based and machine learning methods.

- Neural Network needs much time to train. Powerful hardware is needed
- It is difficulty to find the optimal architecture for a Neural Network.
- It is almost impossible to verify that a drug pair labelled with zero is a non-interacting pair or interacting pair that is not discovered yet. The procedure cannot identify TN and FP pairs correctly.