scientific reports

Check for updates

OPEN

PS3N: leveraging protein sequence-structure similarity for novel drug-drug interaction discovery

Saminur Islam[1]✉, Ahmed Abbasi[2], Nitin Agarwal[3], Wanhong Zheng[4], Gianfranco Doretto[5] & Donald A. Adjeroh[5]

Adverse drug events represent a key challenge in public health, especially concerning drug safety profiling and drug surveillance. Drug-drug interactions represent one of the most popular types of adverse drug events. Most computational approaches to this problem have used different types of drug-related information utilizing different machine-learning algorithms to predict potential drug interactions. In this work, we focus on genetic information about the drugs, particularly the protein sequence and protein structure of protein targets in drug interaction networks, to predict potential drug interactions. We collected various drug information like drug-drug interaction (DDI), Drug attributes like drug active ingredients, protein targets, protein sequence, protein structure etc. We proposed a similarity-based Neural Network framework called protein sequence-structure similarity network (PS3N) and used this to predict novel DDI's. The drug-drug similarities are computed using different categories of drug information based on multiple similarity metrics. Our method outperforms the state-of-the-art and achieves competitive results. Our performance evaluations on different datasets showed the predictive performance as follows: Precision 91%–98%, Recall 90%–96%, F1 Score 86%–95%, Area Under Curve (AUC) 88%–99%, and Accuracy 86%–95%. Our evaluation demonstrates the effectiveness of PS3N in predicting DDI's, including the clinical significance of some new DDI's discovered by the model.

Keywords Protien sequence structure similarity network (PS3N), Drug-drug interaction, Deep learning, Similarity network fusion

Given the increasing number of medications consumed concurrently by individuals, it is becoming more important to know more about the drugs we take. With this increased potential for polypharmacy, there is a corresponding increase in the chance of adverse events induced by medications. It has been revealed that drugs may interact with each other when taken together, and unexpected drug-drug interactions (DDIs) may lead to unanticipated adverse drug events[1,2]. The sheer number of people in polypharmacy has made the issue of drug-drug interactions a significant public health problem. So, it increases the necessity for more accurate predictions and preventive measures to mitigate unforeseen adverse drug events.

Recent biomedical advances have produced vast amounts of drug-related data. Various drug knowledge bases like DrugBank have emerged to manage this data, containing genetic sequences, protein structures, side effects, and chemical structures. Researchers have proposed several approaches to predict drug interactions using data from these sources[3–7]. DrugBank is perhaps one of the most credible databases of known DDIs[8–10] and contains information on over 300,000 DDIs. However, the number of drug-drug interactions is less than 1% of the total possible drug pairs in DrugBank. Thus, many potential interactions remain unidentified, creating gaps in clinical knowledge and decision-making. Additionally, existing machine learning-based methods[8,10–16]

[1]Department of Computer Science, North Carolina State University, Raleigh, NC, USA. [2]Mandoza College of Business, University of Notre Dame, Notre Dame, IN, USA. [3]Department of Information Science, University of Arkansas at Little Rock, Little Rock, AR, USA. [4]Behavioral Medicine and Psychiatry, West Virginia University School of Medicine, Morgantown, WV, USA. [5]Computer Science & Electrical Engineering, West Virginia University, Morgantown, WV, USA. [6]Please note: Abbreviations should be introduced at the first mention in the main text – no abbreviations lists. Suggested structure of main text (not enforced) is provided below. ✉email: sislam8@ncsu.edu

have become popular for predicting DDIs due to their efficiency. These methods are categorized into similarity-based, network-based, matrix factorization-based, and ensemble learning-based approaches[17]. Similarity-based methods use similarities in drug features, such as chemical structures, side effects, or molecular interactions, to predict potential DDIs. Heterogeneous information model like HAN-DDI[18], utilize multiple source of information in Graph Attention Network. However it overlooked the critical importance of protein sequence and structure information, which directly determine the binding affinity, specificity, and mechanism of action, fundamental to understand how drugs interact at the molecular level, unlike broader contextual data like pathways and side-effects. Models like the XGBoost Classifier[19] and MDF-SA-DDI[20] apply similarity metrics such as the Jaccard index and integrate multiple drug features for prediction. Network-based approaches utilize graph-based models to analyze drug networks. Examples include SSI-DDI[21] and MRCGNN[22], which construct molecular graphs and apply graph neural networks (GNNs) to learn drug representations and interactions. Matrix Factorization-based Approaches such as MRMF[23], DDINMF[23], and TMFUF[24] approximate drug interaction matrices using factorization techniques to reveal latent relationships between drugs based on known interactions, side effects, and other features. Ensemble Learning-based Approaches integrate predictions from multiple models, combining different classifiers or feature representations to improve accuracy. For instance, the meta-learning approach[25] combines base classifiers trained on various drug features, and MCFF-MTDDI[26] uses multichannel feature fusion for multiple DDI types. However, Current DDI prediction methods, like similarity-based, network-based, and matrix factorization approaches, focus on surface-level features (chemical structures, side effects) which may overlook more subtle functional interactions. Moreover, Network-based and matrix factorization models often fail to capture novel or mechanistic insights that are critical for identifying previously unknown DDIs. Ensemble learning approaches improve accuracy but are still limited by the quality and depth of the features they integrate.

DDIs are a leading cause of Adverse Drug Events (ADEs), contributing significantly to the global healthcare burden[27–29]. Traditional methods of identifying DDIs during clinical trials are hindered by the complexity of drug interactions, which depend on factors like dosage, genetic variability, and demographic characteristics. Given the limitations of clinical trials in detecting rare or long-term DDIs, there is a pressing need for predictive computational models that can identify potential DDIs earlier in the drug development process.

Existing computational models, including those that utilize social media data, have made some progress in detecting DDIs through indirect signals such as patient reports of adverse events[30]. However, these models often lack the mechanistic depth needed to explain the biological basis of the interactions. Moreover, relying on social media data introduces biases related to reporting accuracy and demographic representation. Our approach diverges from these by integrating rich biological data–namely, protein sequence and structure information–into DDI prediction models.

A drug interaction can occur when two (or more) drugs interact or when a drug interacts with food, beverage, or supplement. Drug interactions can reduce the effectiveness of medication, induce unanticipated side effects, or boost the impact of a drug. Some drug interactions can be dangerous or even life-threatening. Therefore it is always recommended for patients to read the label each time taking a medication, to know more about the potential drug interactions[31]. Fortunately, clinically significant interactions are often predictable and usually undesired (see Some Drugs With Potentially Serious Drug-Drug Interactions)[30]. However, clinicians often find it challenging to rely on predictable drug-drug interactions to achieve the intended therapeutic effect[32]. It was reported that drug interaction is a leading cause of adverse drug events and a significant obstacle for current clinical practice[30].It has led to a considerable public health burden. In the United States alone, > 500,000 serious ADEs were reported annually to the US Food and Drug Administration (FDA) during the past five years[33]. Here we reviewed the recent developments in addressing the challenge of ADEs within the range of different domains. We also constructed a set of data sets to search for new DDI's.

While many studies have incorporated protein or protein target information into DDI models, they lacked of directly mining the rich, underlying biology encoded in protein sequences and structures. For example, ISCMF[34] fused target-based similarity scores into a matrix factorization framework but relied solely on predefined drug-target similarity matrices without ever inspecting the amino-acid sequences or three dimensional folds of those targets. DMFDDI[35] likewise encoded drug-target, enzyme, and transporter information as sparse binary vectors (reduced via Princple Component Analysis (PCA)) before fusing them into deep multimodal network. MUFFIN[36] leveraged a biomedical knowledge graph (including drug-target edges) and molecular graphs via message-passing networks but again no uses of genomic information of proteins. Earlier, Kernel-based SVM methods[37] have built drug-protein interaction predictor using Smith-Waterman alignment scores on protein sequences but those were aimed at Drug target interaction prediction rather than modeling DDI through shared targets. And very recent approaches like LAMFP[38] focus exclusively on chemical fingerprint similarities to address cold-start in DDI prediction, omitting protein information entirely. Much earlier, Yildirim et al[39] used a network of drug protein targets and their protein-protein interactions to study drug adverse effects, but no one has used information from the actual raw protein sequences and the 3D protein structures for the purpose of drug-drug interaction prediction.

In contrast, our work is the first to (1) directly embed both protein sequence and 3D-structure representations into the DDI prediction pipeline, (2) compute multiple, complementary similarity metrics focusing on the functional and structural aspects of proteins, which are often overlooked in favor of interaction networks alone, (3) integrate them end-to-end within a deep neural architecture that jointly learns which biological dimensions most powerfully signal interaction risk. By moving beyond proxy features or black-box knowledge-graph edges, our method captures the functional and structural subtleties of drug targets themselves–improving both predictive accuracy and biological explainability of DDI prediction.

The core challenge lies in developing a predictive framework that accounts for the complex molecular mechanisms underlying drug interactions. Many current models use limited data types, ignoring the potential

impact of a drug's interaction with proteins at the structural level. This study aims to fill that gap by using protein sequence and structure information directly to predict DDIs, offering a more granular understanding of drug interactions at the molecular level.

Our contribution is twofold and offers significant advancement in the field of drug-drug interaction (DDI) prediction. **First**, we introduce a novel deep neural network framework that directly integrates protein sequence and protein structure data, which differentiates it from previous approaches that rely primarily on PPI networks, side effect data, drug interaction networks or indirect sources such as social media. This direct use of biological data provides a more nuanced understanding of the molecular mechanisms behind drug interactions, enabling our model to capture subtle interactions that might be overlooked by traditional methods. **Second**, our model enhances the prediction of adverse drug events by uncovering novel DDIs that may not be detectable through conventional approaches. By leveraging richer, biologically grounded data sources, we provide a more comprehensive and accurate method for predicting DDIs. Beyond improving prediction accuracy, we also explore how these newly predicted interactions can contribute to adverse events, offering a clinically meaningful framework to quantify and assess these risks. This dual focus–on both innovation in prediction and its practical clinical implications–demonstrates the transformative potential of our approach, pushing the boundaries of what is possible in DDI prediction and prevention.

The remainder of this paper is organized as follows. In Section II, we review the existing literature on DDI prediction, highlighting the gaps our research aims to address. Section III details the methodology of our proposed approach, explaining how we integrate protein sequence and structure data into a deep learning framework. Section IV presents the experimental results, comparing the performance of our model with state-of-the-art methods. In Section V, we discuss the broader implications of our findings, including limitations and areas for future improvement. Finally, Section VI concludes the paper with directions for future research.

Related work

Most existing approaches for DDI prediction are based on different properties of the drug compound, such as its chemical structure, side effects, drug-target relationship, and many more. DDIs can be identified with in vivo models using high-throughput screening[40]. However, the price of such procedures is relatively high, and testing large numbers of drug combinations is not practical[41]. To reduce the number of possible drug combinations, numerous computational approaches have been proposed[3,4,4–6,42,43]. In some of these computational approaches, drug-target networks are constructed, and DDIs are detected by measuring the strength of network connections[43], or by identifying drug pairs that share drug targets or drug pathways, for instance, using the random walk algorithm[4].

Some computational approaches have used drug pairs' structural similarity and side effect similarities. For example, Gottlieb et al. proposed the Inferring Drug Interactions (INDI) method, which predicts novel DDIs from chemical and side effect similarities of known DDIs[3]. Vilar et al. used similarities of fingerprints, target genes, and side effects of drug pairs[5]. Cheng et al. constructed features from the Simplified Molecular-Input Line-Entry System (SMILES) data and side effect similarity of drug pairs and applied support vector machines to predict DDIs[42]. Zhang et al. constructed a network of drugs based on structural and side effect similarities and used a label propagation algorithm to identify DDIs[43]. Recently, Ryu et al. proposed DeepDDI, a computational framework that calculates structural similarity profiles (SSP) of DDIs, reduces features using principal component analysis (PCA), and feeds them to a feed-forward deep neural network[7]. The platform generated 86 labeled pharmacological DDI effects, so DeepDDI[44] is a multi-classification (multi-label classification) model.

Some machine learning–based methods have been applied applied to the problem of DDI detection, including approaches based KNN[45], SVM[45], logistic regression[3,42,46] decision tree[42], naive Bayes[42], and network-based label propagation[43] and random walk[47] or matrix factorization[48]. These methods are typically based based on drug properties, such as chemical structure[3,42,43,45,47], Anatomical Therapeutic Chemical classification (ATC) codes[3,42,45], and side effects[3,47,48].

A model was developed to predict DDIs based on the Interaction Profile Fingerprint (IPF)[5]. Quite simply, the interaction probability matrix was computed by multiplying the DDI matrix by the IPF matrix. Afterward,[49] proposed a computational framework by applying matrix perturbation based on the hypothesis that by randomly removing edges from the DDI network, the eigenvectors of the network's adjacency matrix should not change significantly. These two methods employ no other data about drugs except known DDIs.

A new family of similarity-driven methods has followed the assumption that similar drugs should have almost similar interactions. Vilar et al.[49] presented a neighbor recommender method by utilizing substructure similarity of drugs. Relying on Vilar's framework, Zhang et al. constructed a weighted similarity network labeled based on interaction with each of the drugs[43]. They applied an integrative label propagation method using a random walk model on the network to estimate potential DDIs. This prediction framework only considered three types of similarities for predicting DDI via label propagation: substructure-based, side effect-based, and offside effectbased label propagation models[43]. Some methods have also been proposed for adverse event detection using signals from social media[50–52].

In recent years, deep learning is becoming a promising technique for automatically capturing chemical compound features from data sets, and it successfully improves predictive performance. For example, Harada et al.[53] constructed a dual graph convolutional neural network to predict DDIs by combining drugs' internal and external graph structures with learning low-dimensional representations of compounds. However, this method works well only for moderately dense chemical networks with heavy-tailed degree distributions. Tanvir et.al.[18] introduces HAN-DDI, a novel Heterogeneous Graph Attention Netowkr designed to predict -drug-drug interactions (DDIs). The model uses information like drug-protein, drug-pathway, and drug-indication interactions alongside the chemical structure, side-effects and ATC code of drugs. Wang et al.[54] combined interview information on drug molecular and intra-view of DDI relationships, developing a graph contrastive

learning framework to predict DDIs. Lin et al.[55] merged several data sets into a vast knowledge graph with 1.2 billion triples, constructing a large knowledge graph neural network (KGNN) to resolve the DDI prediction. On the other side, based on the structural, gene ontology term, and target gene similarity profiles, Lee et al.[44] applied an autoencoder to reduce the dimensions of each profile, constructing a DNN model by combining all the reduced features to predict the types of DDIs. Deng et al.[56] used the chemical substructures, targets, enzymes, and pathways of drugs to compute a similarity matrix of drugs, inputting each matrix to a DNN model and combining the four submodels to predict DDI events. Besides DDI prediction, deep learning is also successfully applied for drug–target interaction prediction; for example, Shang et al.[57] develop a multilayer network representation learning method to learn the feature vectors of drugs and targets. An and Yu[58] use biased Random Walk and Word2vec algorithms to obtain the feature representation of drugs and targets[59]. Faizan et al[52] introduced DeepSAVE, a deep learning based approach for adverse event detection using social media data.

In recent works, researchers have emphasized the importance of evaluation strategies that prevent information leakage in DDI prediction and related molecular tasks[21]. For example Lv et al.[60] proposed a 3D graph neural network with few-shot learning (Meta3D-DDI) to predict DDI events under a scaffold-based cold start scenario, wherein no drug scaffold (core chemical structure) is shared between training and test sets. This approach addresses the overly optimistic results caused by random splits by ensuring the model is evaluated on structurally novel compounds. Lv et al.[61,62] later extended the scaffold split evaluation to generalize to new chemical scaffolds and thus to a new areas of chemical space. Notably, all these approaches center on small-molecule structure information (2D molecular graph or 3D conformers).

In this study, we develop a novel DDI prediction method utilizing the protein sequence data from the DrugBank[9] and protein structure data from the Protein Data Bank[63]. We calculate different similarity measures to create the similarity matrices for each feature attribute. Then, we use the generated feature matrices to create a single network fusion to measure the potential for interaction between two drugs. The final decision is performed with the help of a neural network architecture based on multilayer perceptrons.

The main novelty of our approach is the focus on only genetic materials (protein sequence and protein structures) associated with the drug targets in developing our prediction model. By integrating protein target information on sequence and structure levels captures the orthogonal aspect of drug relationships that purely chemical structure-based or other drug information based models might overlook. To our knowledge, this is the first attempt at investigating potential DDI prediction by utilizing only information about the protein sequence and structure to generate the feature space fed to the neural network. This biologically informed perspective differentiates our work from prior DDI methods and scaffold-splitting evaluations, while still aligning with the shared goal of improving generalization to novel drugs.

Methodology

We developed a novel neural network model for the prediction of DDIs. The key idea in our approach is that if two drugs have a similar pattern of similarity with other medications, they are likely to have a similar pattern of interacting partners. To capture the patterns of similarity between drugs, we use information about the protein sequences and structures associated with the protein targets for a given drug. Thus, we construct similarity matrices between drugs based on the protein sequences and protein secondary structures and combine these into one protein sequence-structure similarity matrix using network fusion. Figure 1 (adapted from Islam et.al.[64]) shows a schematic diagram of the general proposed framework. To calculate the similarity matrices, we have used cosine distance, Levenshtein distance, Jensen Shannon (JS) divergence, and Euclidean Distance as the similarity measures between a pair of drugs.

Preliminaries

Before discussing the process of predicting drug-drug interaction, we will discuss some primary and essential terminologies and key concepts needed to consider the problem being addressed. All the preliminary knowledge for the methodology will be found in Supplementary (S1. Methodology Preliminaries). You provided information on different distance metrics and how those metrics have been used to calculate the similarity networks from the given protein sequence and protein structure information.

Protein drug targets

Protein drug targets are proteins found in living animals that are linked to specific disorders for which medications are typically used to achieve the desired therapeutic effect. As a result, the protein must be connected to a disease process to be a Protein drug target. Protein drug targets include enzymes, receptors, and transporter proteins. However, receptors account for the majority of the targets. The final number of proteins determined to be the target of an approved small molecule drug was 1324, of which 1249 were found in DrugBank and 313 in the Therapeutic Target Database (TTD). 238 of the proteins were common to both sources, while 1011 were unique to DrugBank and 75 were unique to the TTD[65].

Drug action

The molecular physiological mechanisms by which a chemical creates a response in living organisms are known as drug action. The alterations we observed after taking medications are referred to as drug action effects. Penicillin, for example, interferes with bacterial cell wall formation, resulting in the bacteria's death. Drugs are mainly utilized to distinguish between normal metabolic processes and any anomalies. Because the differences may not be significant, drugs may work in a non-specific manner, altering both normal and unwanted processes–these results in unfavorable side effects.

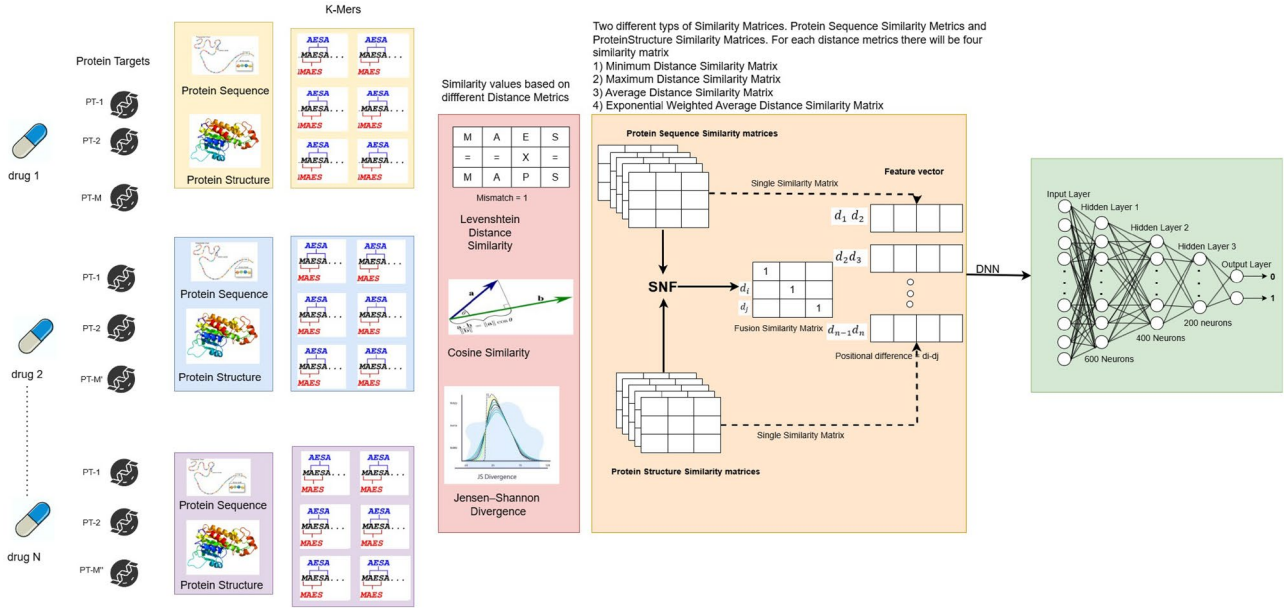

Fig. 1. Proposed protein sequence-structure similarity network (PS3N) model for predicting adverse drug events. Using the method of Similarity Network Fusion (SNF) we create a single $N \times N$ fusion matrix for N drugs. From the fusion matrix, we compute the feature vectors for each pair of drugs. PT denotes protein targets. In this way, we will have possible $\binom{N}{2}$ rows and each row will have N columns as features. These feature vectors are then fed into a multi-layer perception model. For the protein sequence similarity network, the number of hidden layers would reduce to 3 since we have less number of drugs.

Drug pairwise similarity vs drug action similarity

Drug pairwise similarity is a mathematical representation of a relationship between two drugs based on their protein information, which can be either sequence or structural information, with which two drugs would be similar if they were both used for the same ailment. We can construct similarity measures between two drugs using several levels of information such as chemical structure, drug targets, side effects, indications, routes, and so on. We employed drug-protein targets to calculate drug pairwise similarity in this paper.

On the other hand, drug action similarity is a measure of how similar two medications are in terms of how they affect living creatures. The general drug-cell interaction can be used to determine the similarity of drug activity. If two different drugs have active implications on the same cell, we can say they have similar actions, but that doesn't guarantee they will have the same effect on the human body.

Drug similarity space

Drug Similarity space is a feature space of a list of pairs of drugs with some attributes with which we can decide whether a pair of drugs will interact. If there is N number of drugs, then we could have $N^2 - N$ possible pairs that may or may not interact.

K-mer

A k-mer is just a sequence of k characters in a string (or nucleotides in a DNA sequence). It is important to remember that to get all k-mers from a sequence, you need to get the first k characters, then move just a single character for the start of the next k-mer, and so on. Effectively, this will create sequences that overlap in k-1 positions. Decomposing a sequence into its k-mers allows this set of fixed-size chunks to be analyzed rather than the sequence, which can be more efficient. K-mers are very useful in sequence matching (string matching with n-grams has a rich history), set operations are faster and more accessible, and there are a lot of readily available algorithms and techniques to work with them[66].

In our work, we use $k = 4$ for k-mer decomposition, meaning each sequence is broken into overlapping substrings of length 4. From prior research, Yao et al.[67] showed that smaller k values such as 3 and 4, can effectively capture essential sequence information while maintaining manageable feature dimensions. In KAAmer[68], the study demonstrated that using k-mers of length 4 provides a balance between sensitivity and computational efficiency in protein sequence analysis. Furthermore, Zhang et al.[69] proposed a k-mer natural vector method for characterizing protein sequences, considering the numbers and distributions of k-mers including 4-mers for its effectiveness in capturing phylogenetic signals.

Previous research shows that k-mer size can influence model performance. While smaller values may capture more general patterns, larger values yield more motifs. Many protein interaction motifs are very short sequences (often only a few amino acids in lengths). In fact, short linear motifs (SLiMs)[70] that mediate protein-protein interactions are typically only about 3-10 amino acid long (on average 6). Therefore, using a small k for the k-mer

lengths is sufficient to capture these local conserved motifs. Prior studies on sequence-based classification have found that using k in the range of 3-6 yields comparable performance for capturing relevant motif information. Nugent et al.[71] reported no significant difference in model accuracy when using k=3,4,5, or 6 and thus selected k = 4 as an optimal choice for their classification. Consistent with this evidence, we chose k = 4 as it effectively balances motif specificity and coverage: it is long enough to represent the short domain sequence patterns of interest, yet not so long that it over-generalizes to entire domains.

In our study, we did not perform an exhaustive grid search over due to the computational constraints. However, preliminary experimentation with *k* values of 3, 4, 5, and 6 showed stable behavior, and $k = 4$ provided a good trade-off between resolution and representation size for our use case. We acknowledge that further systematic tuning of *k* could offer deeper insight and leave this as a valuable direction for future work.

Suffix array

A suffix array is a sorted array of all suffixes of a string. The definition is similar to a Suffix Tree, a compressed trie of all suffixes of the given text. Any suffix tree-based algorithm can be replaced with an algorithm that uses a suffix array enhanced with additional information and solves the same problem at the same time complexity.

In short, the array of indexes to the sorted array of substrings generated during the transform is essentially a suffix array, which in turn is a representation of the information in a suffix tree[72].

Distance matrices

We compute distance measures (and sometimes similarity measures) between medications based on their protein sequences and structure to estimate the similarity between drugs. In data mining, a similarity measure is a distance with dimensions representing object features. When the distance between two items is small, they are closer to being similar. However, we will see a low degree of resemblance when the distance is significant. There are several different types of similarity or distance metrics. However, in this work, we'll investigate the following: Cosine Similarity, Levenshtein Distance, Jensen Shannon (JS) Divergence, following the methodology described in Islam et al.[64].

Cosine similarity (CS)

Cosine similarity metric finds the normalized dot product of two vectors. By determining the cosine similarity, we would effectively try to find the cosine of the angle between the two objects, when represented as vectors. The cosine of 0° is 1, and it is less than 1 for any other angle. For two *n*-length vectors *A* and *B*, we have:

$$CS(A,B) = \frac{A.B}{\|A\|\,\|B\|} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}} \tag{1}$$

This description is adapted from Islam et al.[64].

Levenshtein Distance (L)

The Levenshtein distance is a string metric for measuring the difference between two sequences. The Levenshtein distance between two strings a, b (of lengths $|a|$ and $|b|$, respectively) is given by $L_{a,b}(|a|, |b|)$

$$L_{a,b}(i,j) = \begin{cases} max(i,j) & \text{if min(i,j)} = 0 \\ min\begin{cases} L(i-1,j)+1 \\ L(i,j-1)+1 \\ L(i-1,j-1)+1 \end{cases} & \text{otherwise} \end{cases} \tag{2}$$

Essentially, $L_{a,b}(i,j)$ is the distance between the first *i* characters of *a* and the first *j* characters of *b*.

This method is discussed in detail in Islam et al.[64].

Jensen Shannon (JS) divergence(JSD)

The Jensen–Shannon divergence is a method of measuring the similarity between two probability distributions. Given two distributions *X* and *Y*, the JS divergence is the average Kullback-Leibler (KL) divergence of *X* and *Y* from their mixture distribution, *M*:

$$JS(X||Y) = \frac{1}{2}D(X||M) + \frac{1}{2}D(Y||M) \tag{3}$$

where $M = \frac{X+Y}{2}$. and *D*(*X*||*M*) is the KL divergence between *X* and *M*. This approach follows the methodology in Islam et al.[64].

Similarity matrices

In most of the DDI prediction methods, it takes a lot of work to find and develop the computational approaches that are appropriate for drug features. For this study, we consider multiple data sources to collect different types of drug feature information to calculate the similarity matrices. In our work, we use similarity matrices rather than distance matrices. Thus, we convert the values into similarity measurements for each distance measure. Here, a detailed discussion is provided for the types of similarity metrics used in our evaluation and the methodology for their computation.

Protein sequence and structure similarity matrices
Each protein structure could have multiple chains. Moreover, each drug's active ingredient could have various protein targets. Thus, we could compute the similarity between two drugs (or drug active ingredients) based on the protein chains associated with the respective protein targets for the drugs. For each similarity measure, we record 1) Minimum Similarity 2) Maximum Similarity 3) Average Similarity(AS) 4) Exponential Weighted Average Similarity (EWAS) as described in Islam et al.[64]. Here, we discuss briefly the protein sequence and protein structure similarity matrices used in this work.

Protein sequence similarity matrices
To compute protein sequence similarity, we utilize several metrics, such as Levenshtein distance, cosine similarity, and Jensen-Shannon (JS) divergence. These metrics are calculated based on *k*-mer profiles derived from protein sequences. The process involves breaking down each protein sequence into subsequences of length *k* (*k*-mers), and constructing *k*-mer profiles for each sequence using the suffix array data structure[72]. The similarity between two proteins is then computed using the following approaches:

- *Levenshtein distance:* This measures the number of single-character edits needed to transform one sequence into another. We normalize the distance and convert it into a similarity score by subtracting the normalized distance from 1.
- *Cosine similarity:* The *k*-mer profiles are treated as vectors and th cosine of the angle between them is calculated to determine the similarity.
- *JS Divergence:* We compute this metric to measure the similarity between two probability distributions, which are derived from the *k*-mer profiles.

Protein structure similarity matrices
Protein structures are more complex, consisting of multiple chains, each with its own 3D configuration. To simplify the comparison of protein structures, we convert each protein's 3D structure into a string-based representation called **pString**, following the method in[73]. This transformation allows us to treat the protein structure as a sequence and apply similarity measures in a similar fashion to protein sequences.

We formalized the similarity calculation to quantify the similarity values between two drug-active ingredients (DAIs). Each DAI can interact with multiple protein targets, and while each protein target has a single sequence, it may consist of multiple chains in its 3D structure. To capture the protein structure information for a given DAI, we represent it as follows:

$$\left[r_1^1, r_2^1, \ldots, r_{k_1}^1; r_1^2, r_2^2, \ldots, r_{k_2}^2; \ldots; r_1^M, r_2^M, \ldots, r_{k_M}^M\right]$$

where r_i^j denotes the *i*-th chain of the *j*-th protein targets, and *M* is the total number of protein targets associated with the DAI, and $k_1, k_2 \ldots k_M$ indicate the number of chains corresponding to each protein target.

This structured approach enables us to systematically compute similarity between the protein structures of different DAIs. For example, If two DAIs have *N* and *M* protein targets with $k_1, k_2, \ldots, k_N$ and $l_1, l_2, \ldots, l_M$ chains respectively, the number of pairwise comparisons (P_c) between the chains of the two DAIs is:

$$P_c = \sum_{i=1}^{N} \sum_{j=1}^{M} k_i l_j \tag{4}$$

This results in a vector of similarity values between the two DAIs, which we use to calculate the minimum, maximum, average, and exponentially weighted average similarity (EWAS) following Islam et al.[64]. The EWAS emphasizes higher similarity values, which we hypothesize are more biologically significant. The weight (w_i) for each similarity value (s_i) is computed as:

$$w_i = \frac{\exp(s_i)}{\sum_i \exp(s_i)} \tag{5}$$

The weighted average similarity is then calculated as:

$$w_{\text{avg}} = \sum_{i=1}^{L} w_i s_i \tag{6}$$

where *L* is the number of chain pairs compared between two protein structures ($L = M \times N$). Additionally, we compute minimum similarity as min_s, maximum similarity as max_s and average similarity as a_s:

$$min_s = \min\{s_1, s_2, \ldots, s_L\}; \qquad max_s = \max\{s_1, s_2, \ldots, s_L\}; \qquad a_s = \frac{\sum_{i=1}^{L} s_i}{L} \tag{7}$$

These statistics allow us to capture different levels of similarity for further analysis. We choose these similarity measures based on their ability to effectively capture both sequence and structural variations. We consider **Levenshtein Distance** because this can captures the mutational and deletional differences between protein sequences. **Cosine Similarity** commonly applied in sequence and text analysis, which provides an efficient way

to compare k-mer profiles in a high-dimensional vector space. Similarly **JS Divergence** compare probability distributions derived from k-mer profiles. Finally **EWAS** emphasizes biologically relevant, higher similarity scores, which are crucial for accurate DDI prediction. The goal was to use these metrics to provide a robust and comprehensive similarity assessment between drug pairs based on protein features.

Protein sequence-structure similarity network (PS3N)

As illustrated in Fig. 1, the left side of the diagram depicts the process of generating similarity metrics. Using protein sequence and protein structure similarity matrices, we construct protein sequence-based and protein structure-based similarity networks through the Similarity Network Fusion (SNF) approach. Each network can be analyzed independently to explore potential drug-drug interactions (DDIs) between drugs or drug-active ingredients[64]. To enhance overall predictive performance, we subsequently integrate the sequence-based and structure-based similarity networks into a comprehensive network, resulting in the Protein Sequence-Structure Similarity Network (PS3N). Our integration process is grounded in the SNF technique, which effectively combines multiple data sources into a single graph that captures relationships among samples[74].

During the similarity network construction and fusion process, we employ the k-nearest neighbors (KNN) approach with $k = 5$ to construct sparse similarity matrices for each type of similarity matrix based on different distance measures. In this approach for each drug, only its top-K most similar neighbors are retained in the matrix, with all other similarities set to ZERO. This sparsification strategy emphasizes the most locally reliable associations, helping reduce the influence of noise and spurious global similarities.

This local filtering aligns with the underlying design of the Similarity Network Fusion (SNF) algorithm[74], which assumes that high local similarity values tend to be more trustworthy than weaker or distant ones. The SNF algorithm iteratively propagates similarity information across networks, such that even if a particular pair of drugs is not among each others top-K neighbors in one modality, consistent associations across different data types can still be reinforced over time through fusion process. Thus, SNF enables the retention of weak but consistent biological signals while dampening modality-specific noise.

Our choice of $k = 5$ follows presedent set by prior work in drug-drug interaction prediction using SNF, particularly the NDD model[75], which also fixed $k = 5$ for both KNN based imputation and SNF sparsification steps. In their study, the authors linked the k used for preprocessing (imputation) to the SNF fusion, nothing that this design avoids introducing additional free parameters. Furthermore, the original SNF paper reports the method to be robust to different values of K, indicating that performance is relatively stable across a reasonable range. based on this evidence and to maintain consistency with validated prior work, we adopted $k = 5$ in our study, balancing robustness, reproducibility, and biological interpretability. The integrated network from SNF serves as the foundation for our analysis of adverse drug events, particularly in the context of predicting DDIs. While our model primarily focuses on protein sequence and structure similarity for DDI prediction, we also incorporate direct DDI pair information during the dataset labeling process. Specifically, we leverage DDI pair data provided by DrugBank to label known interacting drug pairs, which informs our training of the PS3N model.

Neural network model

Our proposed neural network model is tailored specifically to the datasets we are utilizing, meaning its performance is inherently influenced by the diversity and quantity of medications represented within the dataset. The architecture of our neural network consists of four hidden layers, with the number of neurons in each layer fine-tuned through cross-validation to optimize performance as described in Islam et al.[64]. Each hidden layer employs the Rectified Linear Unit (ReLU) activation function, which is defined as follows:

$$f(x) = x^* = \max\{x, 0\} \tag{8}$$

The ReLU activation function is chosen for its computational efficiency and effectiveness in mitigating the vanishing gradient problem, which is particularly beneficial when training deep networks. The final output layer utilizes the sigmoid function to generate interaction probabilities between drug pairs, formulated as:

$$\text{Sigmoid}(x) = \frac{1}{1 + e^{-x}} \tag{9}$$

To avoid overfitting, we integrated dropout layers following each hidden layer, with dropout rates varying between 0.3 and 0.5. This technique randomly ignores a subset of neurons during the training process, which enhances the network's ability to generalize and prevents co-adaptation of neurons, making the model more robust. We employ **Xavier weight initialization** for each layer to ensure that the weights are properly scaled, facilitating faster convergence during training. This initialization strategy is particularly useful in preventing issues related to the gradient propagation in deep networks.

For our loss function, we utilize categorical cross-entropy, which is suitable for multi-class classification. This choice maintains the generality of our predictions, even when the output labels are one-hot encoded, focusing only on the positive class C_p. One element in the target vector t is not zero $t_i = t_p$. So discarding the aspects of the summation that are zero due to target labels[76], we can write Cross-Entropy (CE) as,

$$CE = -log(\frac{exp(S_p)}{\sum_j^C exp(S_j)}) \tag{10}$$

where S_p is the CNN score for the positive class.

Adam Optimizer is used as the optimizer which is particularly well-suited for large, sparse datasets. A learning rate is maintained for each network weight (parameter) and separately adapted as learning folds. It combines momentum and the Root Mean Square propagation (RMSProp) process to speed up the learning process. If m, v represents the momentum vector and β_1, β_2 as the exponential decay, the update rules of Adam would be[77],

$$\hat{m}^{k+1} = \frac{m^{k+1}}{1-\beta_1^{k+1}} \tag{11}$$

$$\hat{v}^{k+1} = \frac{v^{k+1}}{1-\beta_2^{k+1}} \tag{12}$$

which will finally represent the optimization function as

$$\theta^{k+1} = \theta^k - \eta \frac{\hat{m}^{k+1}}{\sqrt{\hat{v}^{k+1}} + \epsilon} \tag{13}$$

Experiments and results

For the development of the PS3N model to predict effectively the DDI's, we put significant time into data processing and cleaning. Our whole experimental setup is divided into multiple parts. Before going to the experiments and results here is a brief introduction to the datasets we utilize to prepare our similarity matrices.

Datasets

In this work, we utilized multiple sources to gather comprehensive drug-related information. Most of the data, including protein sequences and pathways, was obtained from DrugBank (https://go.drugbank.com/). Additionally, protein structural data was retrieved from the Research Collaboratory for Structural Bioinformatics (RCSB) Protein Data Bank (https://www.rcsb.org), where protein chains were extracted from PDB files using Biopython libraries. These sources were combined to create a dataset of 905 drugs (active ingredients) from DrugBank, containing both protein structure and protein sequence information. Information on drug side effects and drug indications was sourced from the Side Effect Resource)(SIDER) database (http://sideeffects.embl.de/), while protein-protein interaction data was collected from the UniProt portal (https://www.uniprot.org/uniprot/).

The DrugBank dataset was primarily used to train and evaluate the performance of the proposed model. Furthermore, we compared the performance of our model on the DrugBank dataset with previously published works that also utilized this dataset for similar analyses. In addition to the DrugBank dataset, we employed two benchmark datasets, DS1 and DS2, originally reported by Rohani et al.[75], to further assess our model's performance and validate its predictive capabilities against state-of-the-art methods in drug-drug interaction prediction.

Performance evaluation of PS3N model

To evaluate the performance of the proposed method, we compared it with machine learning approaches such as kNN (k-Nearest Neighbor), RF (Random Forest), Logistic Regression, LDA (Linear Discriminant Analysis), and Support Vector Machine. We also compared our results with state-of-the-art methods proposed in[75,5,43,78–84].We evaluated the competitiveness of our models using different performance metrics such as Precision, Recall, F1, Area under Curve (AUC), and AUC-PR (AUC using the precision-recall curve).

Here, these are defined as follows:

$$\text{Precision} = \frac{TP}{TP+FP} \qquad \text{Recall} = \frac{TP}{TP+FN} \qquad \text{F-measure} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{14}$$

where TP, TN, FP, and FN stand for True Positive, True Negative, False Positive, and False Negative. Precision is the fraction of correct predicted interactions among all predicted interactions. while recall is the fraction of correct predicted interactions among all true interactions. Precision and recall have a trade-off thus improving one of them may lead to a reduction in another. Therefore, utilizing F-measure which is the geometric mean of precision and recall is more reasonable.

We note that if the interaction between two drugs is assigned to zero, it simply implies that no evidence of their interaction has been found yet. The two may still interact, but the features we have used so far are not able to detect the interaction.

Thus, we cannot identify TN and FP pairs correctly. The training process requires both positive and negative samples. Therefore, some of the zero assigned pairs are considered as non-interacting pairs in the training model. So every method may have some FP in its evaluations. This leads to a reduction in calculated precision and F-measure, while the real values of precision and F-measure may be higher. Since the values of precision, recall, and F-measure are dependent on the value of the threshold, we also evaluate the methods via AUC which is the area under the receiver operating characteristic (ROC) curve.

Experimental setup

For our training experiment, we split each dataset into training, validation, and test sets according to a $70\%/10\%/20\%$ random split to ensure an unbiased representation of data. To avoid any form of data leakage the test set was kept separate and never used for any training, validation, or tuning model parameters. In our

Distance metric	Precision	Recall	F-measure	AUC	Accuracy
L AS	0.9199	0.9419	0.9308	0.9673	0.9081
JSD AS	0.8837	0.8667	0.8751	0.9181	0.8377
CS AS	0.8799	0.9093	0.8943	0.9315	0.8590
L EWAS	0.9499	0.9638	0.9568	0.9832	0.9429
JSD EWAS	0.9406	0.8835	0.9112	0.9559	0.8870
CS EWAS	0.9523	0.9632	0.9578	0.9833	0.9443

Table 1. Performance of PS3N using single similarity matrices based on protein sequences. *L = Levenshtein, JSD = JS Divergence, CS = Cosine, AS = Average Similarity, EWAS = Exponential Weighted Average Similarity

Distance metric	Precision	Recall	F-measure	AUC	Accuracy
JSD AS	0.8796	0.9279	0.90313	0.9171	0.8564
CS AS	0.9037	0.9469	0.9248	0.9499	0.8889
JSD EWAS	0.9762	0.9650	0.9706	0.9895	0.9578
CS EWAS	0.9743	0.9738	0.9741	0.9910	0.9627

Table 2. Performance of PS3N using single similarity matrices based on protein structure. *JSD = JS Divergence, CS = Cosine, AS = Average Similarity, EWAS = Exponential Weighted Average Similarity

study, we conducted a comprehensive analysis of the two important hyperparameters: the optimizer and learning rate. These experiments were designed to assess the impact of different combinations of these parameters on the model's accuracy, a key metric for the effectiveness of the PS3N model. For each dataset, networks were trained on the training set for a total of 100 epochs with a batch size of 100 for the proposed neural network method.

We initialize the weights of the network so that the neuron activation function can avoid saturation problems or be stuck in dead regions. We used a batch size of 100 and 20 - 50 epochs, with Categorical cross entropy and Adam Optimizer for optimization with a momentum parameter of 0.9. The number of epochs was set to 20 and 50.

To provide a comprehensive evaluation, our experiments include both a standard random split for performance benchmarking and a challenging inductive split (cold-start split) for generalization to strike a balance between comparability with prior work and rigor in testing PS3N's ability to handle novel cases.

Evaluation of single similarity metrics for PS3N model

To evaluate the contribution of individual similarity metrics to the performance of the PS3N model, we tested the model using single feature matrices derived from protein sequence and protein structure. The similarity matrices were computed using various distance metrics, including Levenshtein (L), JS Divergence (JSD), and Cosine Similarity (CS), under two weighting schemes: Average Similarity (AS) and Exponential Weighted Average Similarity (EWAS).

Tables 1 and 2 present the result of these evaluations. Table 1 focuses on protein sequence-based simialrity matrices, while Table 2 explores protein structure-based similarity matrices. From these results, several key observations emerge. Firstly, The exponential weighted average similarity (EWAS) consistently outperforms the simple average similarity (AS) for both protein sequences and protein structures. For example, CS EWAS achieves an F-measure of 0.9578 and an AUC of 0.9833 for protein sequences, significantly higher than CS AS (F-measure of 0.8943, AUC of 0.9315). Protein structure-based similarity matrices using EWAS metrics show even higher performance. Specifically, CS EWAS achieves an F-measure of 0.9741 and an AUC of 0.9910.

Secondly, Protein structure-based metrics demonstrate slightly higher predictive power compared to sequence-based metrics. This highlights the importance of integrating structural information in drug-drug interaction (DDI) predictions. The results underscore the importance of evaluating single similarity metrics to understand their contribution to the overall performance. Metrics like CS EWAS for structure and sequence exhibit stability and reliability, making them vital components of the final fused similarity model.

Comparative analysis on drugBank dataset

Table 3 presents the results of the PS3N model when using fused similarity metrics, combining protein sequence and protein structure information, compared with recent state-of-the-art models evaluated on the DrugBank dataset. PS3N achieves strong performance with high precision (0.980) and recall (0.982), alongside competitive AUC (0.995) and accuracy (0.973). It is important to note the limitation in directly comparing precision and recall metrics, as several state-of-the-art models (e.g., DSIL-DDI[83], AutoDDI[84], R2-DDI[81], GMPNN-CS[82], SSI-DDI[21]) reported only aggregated metrics such as F-measure and AUC. Given these reporting limitations, a complete and statistically rigorous comparison across all evaluation metrics on the DrugBank dataset was challenging. However, detailed comparisons with fully reported metrics and statistical analyses are comprehensively provided on benchmark datasets DS1 and DS2.

While models such as DSIL-DDI[83] and AutoDDI[84] achieve marginally higher F-measures, their absence of detailed precision and recall reporting raises concerns about potential trade-offs between minimizing false

Method	Precision	Recall	F-measure	AUC	Accuracy
SVM	0.551	0.208	0.302	0.571	0.732
KNN	0.547	0.619	0.581	0.710	0.751
NDD[75]	0.565	0.193	0.287	0.737	0.731
R2-DDI[81]	N/A	N/A	0.981	**0.997**	**0.982**
GMPNN-CS[82]	N/A	N/A	0.954	0.985	0.953
SSI-DDI[21]	N/A	N/A	0.964	0.989	0.963
DSIL-DDI[83]	N/A	N/A	**0.993**	0.989	0.981
AutoDDI[84]	N/A	N/A	0.976	0.995	0.976
PS3N (ours)	**0.980**	**0.982**	0.981	0.995	0.973

Table 3. Performance comparison of the proposed PS3N model with state-of-the-art methods using 1107 drugs from the DrugBank dataset. Bold values indicate the best performance, and underlined values indicate the second-best performance for each metric.

Method	AUC	AUC-PR	F-measure	Recall	Precision
Substructure-based label propagation model[43]	0.937	0.901	0.804	0.797	0.811
Side-effect-based label propagation model[43]	0.936	0.903	0.806	0.793	0.820
Offside-effect-based label propagation model[43]	0.937	0.904	0.809	0.795	0.823
Vilar's substructure-based model[5]	0.936	0.902	0.804	0.797	0.812
Classifier ensemble method[78]	0.956	0.928	0.836	0.827	0.843
Weighted average ensemble method[78]	0.948	0.919	0.831	0.835	0.826
NDD[75]	0.954	0.922	0.835	0.836	0.833
PS3N (Protein Sequence) (ours)	**0.974**	0.948	0.916	0.925	**0.906**
PS3N (Protein Structure) (ours)	0.972	**0.949**	**0.917**	**0.932**	0.903
PS3N (Sequence + Structure) (ours)	0.972	0.948	**0.917**	0.931	0.903

Table 4. Performance comparison of different methods on DS1 from[75]. We obtained information on 469 drugs for protein sequences, and on 414 drugs for protein structure. Results in the first six rows are taken from[75]. Bold values indicate the best performance for each metric.

positives and maximizing true positives. Similarly, R2-DDI[81] and GMPNN-CS[82] provide strong F-measure and AUC values but lack a precision-recall balance assessment, which is critical for validating their consistency across different datasets. In contrast, PS3N excels in both precision and recall, ensuring robust predictions that reduce false alarms (false positives) while capturing true interactions effectively. SSI-DDI[21] , though competitive in F-measure (0.964) and AUC (0.989), does not offer the transparency provided by PS3N, particularly regarding its ability to handle diverse interaction scenarios.

Furthermore, NDD's[75] performance underscores the limitations of simpler models that do not integrate advanced similarity metrics. Its low F-measure (0.287) and AUC (0.737) demonstrate that without leveraging fused metrics, models struggle to capture the complexity of drug interactions. PS3N's fusion of protein sequence and structure-based metrics addresses this limitation by synergizing the complementary strengths of both data types. The integration of exponential weighted average similarity (EWAS) enhances the model's ability to emphasize critical patterns, resulting in superior predictive power and consistency.

In summary, PS3N offers a distinct advantage by combining state-of-the-art performance with detailed reporting of essential evaluation metrics. Its ability to maintain high precision, recall, and AUC, alongside its robustness in diverse scenarios, underscores its superiority over competing methods. The fusion approach not only improves performance but also ensures stability and interpretability, making PS3N a valuable tool for DDI prediction.

Comparative performance on benchmark dataset DS1 and DS2

In Tables 4 and 5 we present a comprehensive performance comparison of our proposed approach with state-of-the-art algorithms, showing several noteworthy strengths in our methodology.

Our comparative analysis focuses on DS1 and DS2 datasets from[75], which allowed us to generate protein sequence and protein metrics for a subset of drugs. By incorporating the new feature space into our model, we have observed significant enhancements in performance, particularly in terms of AUC, Precision, and Recall.

In Table 4, our PS3N approach stands out as a top-performing method, consistently surpassing other existing techniques. This substantial improvement shows the effectiveness of our proposed approach in accurately predicting drug interactions. It's important to note that the performance of PS3N remains consistently strong across datasets based on sequence, structure, or the combination of both, as demonstrated in Table 4. This robustness highlights our model's strength in handling different types of data sources effectively.

Method	AUC	AUC-PR	F-measure	Recall	Precision
Substructure-based label propagation model[43]	0.788	0.208	0.294	0.537	0.197
Vilar's substructure-based model[5]	0.810	0.244	0.312	0.479	0.232
Classifier ensemble method[78]	0.936	0.487	0.553	0.689	0.462
Weighted average ensemble method[78]	0.646	0.440	0.15	0.226	0.118
NDD[75]	0.994	0.890	0.825	0.804	0.847
BioDKG-DDI[79]	0.967	N/A	0.903	0.918	0.884
DPDDI[80]	0.956	N/A	0.840	0.810	0.754
PS3N (Protein Sequence) (ours)	**0.998**	**0.975**	**0.978**	0.987	**0.972**
PS3N (Protein Structure) (ours)	0.997	**0.975**	**0.978**	**0.992**	0.964
PS3N (Sequence + Structure) (ours)	0.997	0.970	0.977	0.987	0.970

Table 5. Performance comparison of different methods on the DS2 Dataset from[75]. We obtained information on 585 drugs for protein sequences, and on 504 drugs for protein structure. Results in The first five rows are taken from[75]. Bold values indicate the best performance for each metric.

On the DS2 dataset as we see in Table 5, PS3N demonstrates clear superiority over recent methods such as BioDKG-DDI and DPDDI. While BioDKG-DDI leverages a knowledge graph-based approach and achieves a competitive F-measure of 0.903 and recall of 0.918, its lower precision (0.884) compared to PS3N (0.972 for Protein Sequence) indicates a higher tendency for false positives. Similarly, DPDDI, a deep learning-based model, achieves an AUC of 0.956 and an F-measure of 0.840 but lags behind PS3N in precision (0.754), underscoring its reduced reliability. PS3N achieves consistently high precision (0.972 for Sequence, 0.964 for Structure), recall (0.987 for Sequence, 0.992 for Structure), and F-measure (0.978 for both Sequence and Structure), reflecting its robust ability to capture complex drug interaction patterns while minimizing false positives. Furthermore, the integration of protein sequence and structure data in PS3N enhances its predictive accuracy by leveraging complementary insights, offering a more balanced and competitive performance compared to standalone methods.

Interestingly, our PS3N method and the NDD approach produced quite similar results, especially when we looked at performance in terms of AUC. This similarity suggests that both methods are good at understanding complex patterns in how drugs interact with each other. In a way, our model competes well with the established NDD method in predicting these interactions. This also holds in Table 5 which was created from the DS2 dataset.

Inductive (Cold-Start) evaluation

To rigorously assess PS3N's ability to generalize to unseen drugs, we adopt a drug-centric cold-start (inductive) split following the SSI-DDI[21] protocol. The full drug set is randomly divided into "Known" (Gold, 80 %) and "New" (Gnew, 20 %) subsets, ensuring no "New" drug appears in training. At test time, we evaluate on two partitions:

- *CS1(New–New):* Both drugs in each pair belong to Gnew (i.e., neither drug was seen during training).
- *CS2 (Known–New):* One drug in each pair is from Gold (seen in training), and the other is from Gnew (unseen).

To handle the information leakage recent studies utilized two types of data splits to evaluate the models performance. Some studies utilized scaffold-based splitting (separating compounds by core structure)[60], which is a rigorous way to avoid information leakage in approaches that use chemical or molecular structures. While scaffold splits eliminates any shared substructures between training and test sets, we opted for a drug-level inductive split in this work for both practical and conceptual reasons. Since PS3N relies exclusively on protein sequence and structure information–without using any chemical or molecular representations of drugs–the scaffold-based splitting strategy commonly applied in molecular modeling is not applicable in our case. Our features are derived entirely from biological similarity among drug targets, and the cold-start drug-level split ensures that the model encounters novel drugs and their associated protein information for the first time at test time.

Inductive split ensures that the model has no prior access to the interaction behavior or protein associations of those test drugs, directly addressing the risk of information leakage related to drug identity or target features. This strategy evaluates PS3N's ability to generalize to entirely unseen drugs, which aligns with our primary goal of enabling reliable DDI discovery for novel compounds. CS1 represents the fully cold-start scenario–predicting interactions between two entirely novel drugs–while CS2 represents a semi-cold scenario–predicting interactions between one familiar drug and one novel drug. Table 6 compares PS3N against other state-of-the-art baselines under these two partitions.

As illustrated in Table 6, PS3N achieves an accuracy of **70.05%** on CS1, surpassing SSI-DDI's 65.02% by over five percentage points. This improvement highlights PS3N's robust capability in accurately identifying interactions when both drugs involved are entirely unseen during training. While PS3N's AUROC score for CS1 (62.12%) is lower than SSI-DDI's 72.38%, its superior accuracy suggests that the decision threshold chosen by PS3N is better optimized for making correct binary predictions in this challenging fully cold-start scenario. In the CS2 setting, where one drug is familiar and the other is new, PS3N attains an accuracy of 69.03%, slightly

Model	Transductive		CS1 (New–New)		CS2 (Known–New)	
	ACC (%)	AUROC (%)	ACC (%)	AUROC (%)	ACC (%)	AUROC (%)
DeepDDI[85]	93.15	99.76	47.96	88.39	68.69	95.75
MR-GNN[86]	94.06	99.76	49.93	91.70	72.38	97.52
MHCADDI[87]	78.50	86.33	66.53	72.13	72.34	79.22
SSI-DDI[21]	94.47	98.38	65.02	72.38	73.35	80.92
PS3N (Ours)	**97.30**	**99.50**	**70.05**	**62.12**	**69.03**	**67.63**

Table 6. Overall performance comparison: transductive vs. inductive settings. Bold values indicate the best performance for each metric.

below SSI-DDI's 73.35% by about four percentage points. Similarly, its AUROC of 67.63% trails behind SSI-DDI's 80.92%. Despite this, PS3N still demonstrates solid performance in binary classification tasks, maintaining competitiveness even when only partial information is available. However, the lower AUROC indicates PS3N's limitations in perfectly ranking predicted interactions compared to graph-based approaches.

Beyond inductive splits, it is instructive to compare PS3N's performance under both transductive and inductive settings. Table 6 summarizes PS3N alongside SSI-DDI and key baselines across transductive (random-pair) and inductive partitions.

In the transductive (random-pair) setting, PS3N achieves impressive results, with an accuracy of **97.30%** and AUROC of **99.50%**, surpassing SSI-DDI's performance of 94.47% accuracy and 98.38% AUROC. However, this high transductive performance may reflect model memorization or overfitting, since drugs in the test set often share similar patterns or neighborhood characteristics with those in the training set. Thus, evaluating models under inductive scenarios (such as CS1 and CS2) provides a clearer assessment of their real-world predictive power. Even in these more challenging contexts, PS3N remains competitive, achieving 70.05% accuracy on CS1 compared to SSI-DDI's 65.02%, and 69.03% accuracy on DS2 versus SSI-DDI's 73.35%.

A detailed examination of these results yields several insights. First, PS3N's notable accuracy advantage on CS1 highlights its ability to generalize effectively, capturing meaningful signals from similarity-based features and pretrained embeddings even when both drugs are completely new. Although PS3N's lower AUROC score (62.12% compared to SSI-DDI's 72.38%) indicates a weaker separation between positive and negative interactions, its accuracy advantage demonstrates that its decision threshold is better tuned for accurately classifying binary outcomes. This practical strength is particularly valuable for real-world drug interaction screening pipelines, where the priority is correctly flagging potential interactions for experimental validation rather than purely ranking them.

Second, in the DS2 scenario where one drug is familiar and the other is unseen, PS3N achieves an accuracy of 69.03%, reflecting its capacity to leverage known contextual information effectively. Yet, its AUROC (67.63%) trails SSI-DDI's 80.92%. This difference suggests that PS3N might benefit from integrating techniques like graph propagation or topological regularization, as used by SSI-DDI, to better embed novel drugs and enhance interaction ranking. Still, the competitive accuracy underscores that PS3N's existing design, which blends PCA-transformed similarity features with fixed embeddings, already adequately supports accurate binary predictions in semi-cold scenarios.

Third, when comparing transductive and inductive scenarios, it becomes clear that PS3N experiences a more substantial decrease in AUROC–approximately 37.38 percentage points from transductive to CS1 and about 31.87 points to CS2–compared to SSI-DDI's more moderate declines of roughly 26.00 and 17.46 points, respectively. This indicates that graph-regularized methods like SSI-DDI inherently produce more stable interaction scores in cold-start conditions. PS3N's reliance on fixed embeddings leads to noisier similarity representations when encountering entirely novel chemical spaces. Nonetheless, PS3N's ability to achieve comparable or superior accuracy, particularly in fully cold-start conditions (CS1), demonstrates its robustness and practical utility for predicting drug interactions.

Reliability of the PS3N model: performance in datasets

To assess the reliability of the proposed PS3N model in predicting drug-drug interactions (DDIs), paired t-tests were conducted to compare its performance with existing state-of-the-art methods in three data sets: DrugBank, DS1, and DS2. The analysis focused on key performance metrics: AUC, AUC-PR, F-measure, Recall, and Precision. The results of the paired t-test are presented in Table 7, where statistically significant improvements ($p < 0.05$) for the PS3N model.

For the DrugBank dataset, the PS3N model showed highly significant improvements in Precision and Recall (both with $p < 0.05$), confirming the model's strength in accurately classifying interactions. While F-measure also showed a marginal improvement, it was not statistically significant ($p = 0.08169$), suggesting that the model's ability to balance precision and recall was not as enhanced in this dataset compared to others.

In DS1 dataset, the PS3N model demonstrated statistically significant improvements across all evaluated metrics compared to baseline methods. The improvements in Precision, Recall, and F-measure were particularly noteworthy, with p-values well below the 0.05 threshold. These results suggest that PS3N (using both sequence and structure information) showed statistically significant improvements over all baseline methods across all evaluated metrics ($p < 0.05$). These results demonstrate the model's robust performance across diverse evaluation measures.

Metric	DrugBank		DS1		DS2	
	t-stat	*p*-value	*t*-stat	*p*-value	*t*-stat	*p*-value
Precision	− 78.005	**0.00016**	− 4.976	**0.0011**	− 3.134	**0.014**
Recall	− 4.600	**0.04415**	− 5.149	**0.0009**	− 3.143	**0.014**
F-measure	− 2.032	0.08169	− 5.096	**0.0009**	− 3.034	**0.016**
AUC	− 2.034	0.08139	− 4.345	**0.0025**	− 2.379	**0.045**
AUC-PR	NA	NA	− 4.465	**0.0021**	2.878	**0.028**
Accuracy	− 2.087	0.07528	NA	NA	NA	NA

Table 7. Paired t-test results for DrugBank, DS1 and DS2 dataset comparing PS3N (Sequence + Structure) with baseline methods. Statistically significant results ($p < 0.05$) are highlighted. Bold values indicate the best performance for each metric.

On the DS2 dataset, the PS3N model again showed significant improvements over baseline methods in several metrics, including Precision, Recall, F-measure, and AUC-PR ($p < 0.05$). Notably, the improvement in AUC was not statistically significant ($p = 0.066$), suggesting that while the model improved in identifying true positives (high Recall) and reducing false positives (high Precision), its ranking ability, as indicated by AUC, did not show a significant difference compared to the baseline methods.

Our statistical analysis confirms that the PS3N model shows statistically significant improvements over baseline methods in most performance metrics for the DrugBank, DS1 and DS2 data sets. Although AUC improvements were not statistically significant in all cases, particularly in the DS2 dataset, this discrepancy can be attributed to factors such as the imbalance of the dataset and threshold sensitivity. However, the model demonstrated substantial improvements in key metrics such as Precision and Recall, with particularly strong performance in the DS1 and DrugBank datasets. These improvements, along with significant gains in the F measure, highlight the robustness of the PS3N model and its practical relevance for reliable prediction of DDI in diverse datasets.

Predicted DDI network using PS3N

We quantified all the drug-drug interactions using the PS3N model in Fig. 2. As a result, when we anticipate drug interactions between pairs of drugs, the model generates a likelihood of interaction between the drugs, which we refer to as an interaction score, which ranges from 0 to 1. We can also use the model to get the projected labels for the test data. To prepare network data, the drug IDs are employed as network nodes. We considered a pair of drugs to have an interaction if the expected labels from the model were 1. We discovered that the network data contained a wealth of interaction information. We proposed imposing a threshold value of 0.7 on the principal similarity network to eliminate some data from consideration to improve presentation. In Fig. 2, we observed several drug-drug interactions (DDIs), all depicted in the same shade. Our primary objective was to evaluate our model's DDI prediction accuracy. As part of our analysis, we applied a threshold to the network's similarity values. When we examined the predictions with the threshold, we found that 89 percent of the interactions were correctly labeled.

However, upon reviewing the entire network without imposing a strict threshold, we achieved an overall DDI detection accuracy of approximately 96 percent. This indicates that even some estimated similarity values falling below the 0.70 threshold still corresponded to real drug interactions.

Our model excels in predicting drug-drug interactions. The key factor here is the threshold we use. If we set a strict threshold, we catch fewer interactions, but we're more certain they're accurate. When we use a more lenient threshold, we capture a wider range of interactions, even if some have lower similarity scores. This approach helps us uncover potential interactions we might miss with a stricter threshold.

In Fig. 2, we can see that the interaction network reveals distinct clusters. A manual examination of the clusters revealed some interesting observations from the interaction network. Each of the distinct clusters of drug interactions represents different pharmacological classes, each with unique therapeutic indications and mechanisms of action. In the top left cluster, where almost all the drugs interact with the remaining drugs, we found three classes of drugs create subclusters within the cluster. These are the Dopamine Agonists Cluster (DBCAT000607), a class of drugs normally used in the treatment of Parkinson's disease and related conditions; Phenothiazines Cluster (DBCAT000801), a class of drugs used in the treatment of various psychiatric disorders, including schizophrenia and agitation, Antipsychotic Agents Cluster (DBCAT000529), a class of drugs used to control psychotic behavior and alleviate symptoms of schizophrenia and related disorders[9]. From this information, we can deduce very important adverse drug events, for example, DAI DB00714(a dopamine agonist) and DB00433(a phenothiazine, which is also a dopamine antagonist) having interaction, which means that the therapeutic effect of DB00714 can be decreased when used in combination with DB00433 because of counteracting effect of these two medications on dopamine receptors[9].

We also found some major subclusters in the top right cluster, which are the Diuretics Drug Cluster (DBCAT00542), a class of drugs to help reduce fluid buildup in the body by increasing urine output, Factor Xa Inhibitors Cluster (DBCAT001775) that inhibit or block Factor Xa activity, thereby exerting antithrombotic effects, Anticoagulants Cluster (DBCAT000007 / DBCAT003243), a class of drugs that prevent clotting, Blood Coagulation Factors Cluster (DBCAT000007 / DBCAT003243), a class of drugs that promote blood clot formation[9]. We found that DDI (DB00606 – DB00562) in DBCAT00542, where the efficacy of DB00606 can

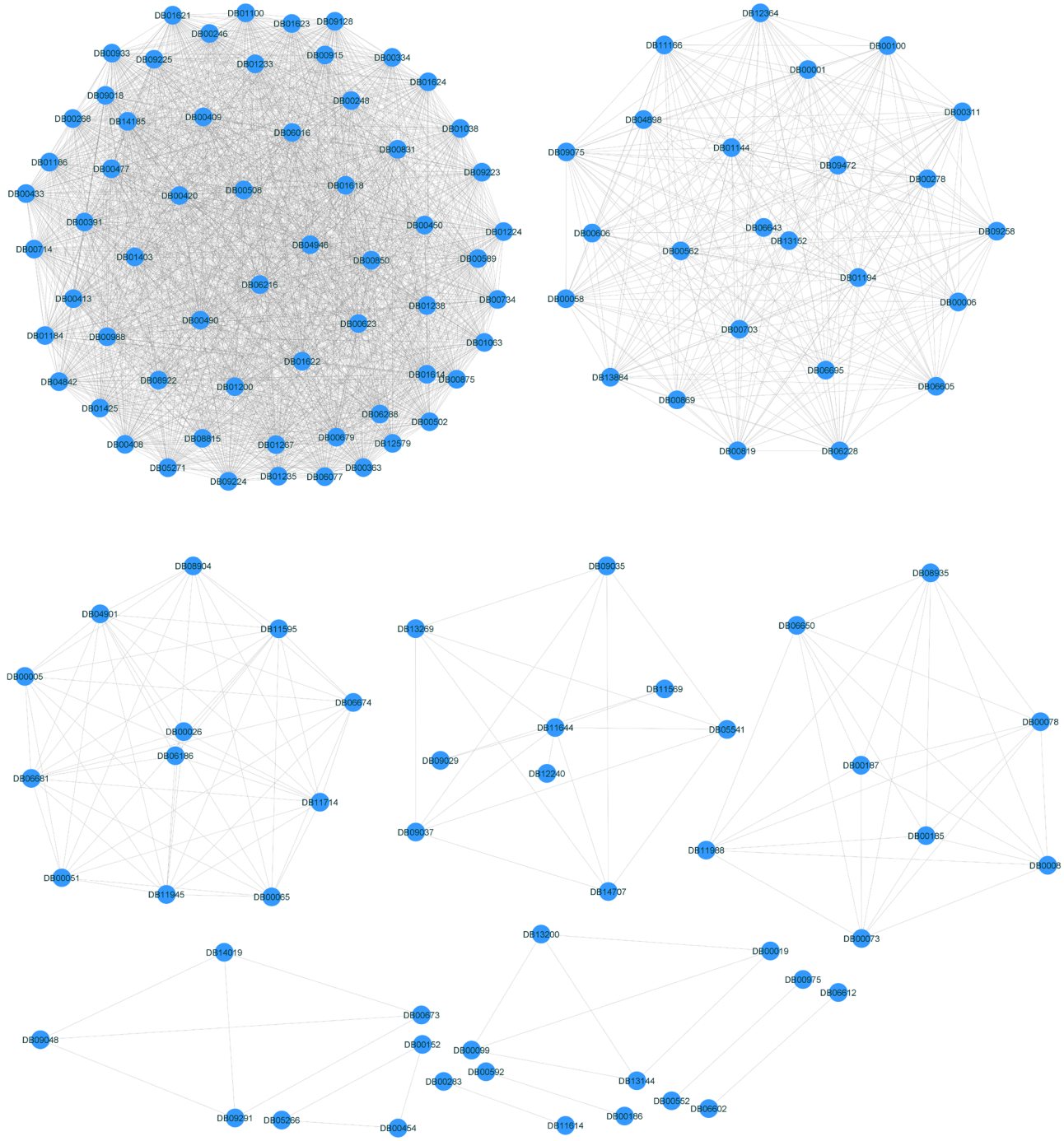

Fig. 2. DDI network diagram constructed from the DrugBank dataset. Interactions with predicted probabilities below 0.7 were excluded, and edges were included only for pairs with a predicted label of 1, resulting in the PS3N DDI network.

be increased when combined with DB00562. On the other hand, DDI (DB00606-DB00703) and DB00703 may increase the hypotensive activities of DB00606[9].

Drug Interaction could be possible between drugs that are not in similar drug classes. For example, DB00606 is interacting with DB06228. The ATC classification of DB00606 is C-Cardiovascular system CO3-DIURETICS[9] on the other hand DB06228 is a class of Xa inhibitors drug. DB00606 may increase the excretion rate of DB06228 which could result in a lower serum level and potentially a reduction in efficacy[9]. This highlights the complexity of drug interactions, as they can occur between medications of different classes and can significantly impact therapeutic outcomes. So, Fig. 2 explains how a similarity network could be useful for finding adverse drug effects and interpreting the positive and negative interactions. Similarity within clusters suggests shared pharmacological properties, while drugs in multiple clusters highlight their diverse pharmacological effects.

Understanding these interaction networks is crucial for predicting potential drug-drug interactions, optimizing therapy, and minimizing the risk of adverse reactions in clinical practice.

Clinical implications of predicted drug interactions

Our model PS3N showed excellent performance in performed very well in predicting the existing drug interactions found from DrugBank. In Fig. 2, we observe the first cluster (top left) in which all the drugs interact with almost all other drugs. This can be explained from a clinical perspective. Dopamine antagonist phenomena are based on the different drug classes. In the cluster, there are multiple drug classes such as Dopamine (DBCAT000607) and phenothiazines (DBCAT000801) as pharmacologically, they all target dopamine receptors and Antipsychotic Agents (DBCAT000529) which are also Dopamine blockage agents. The interactions among the drug classes have significant clinical implications.

The subset of drugs in the top left cluster belongs to the Phenothiazines group, which is a subset of the broader category of Antipsychotic Agents. These drugs include Prochlorperazine (DB00433), Mesoridazine (DB00933), Pipotiazine (DB01621), Thioridazine (DB00679), Acetophenazine (DB01063), Fluphenazine (DB00875), Trifluoperazine (DB00831), Promazine (DB00420), Chlorpromazine (DB00477), Methotrimeprazine (DB01403), and Perphenazine (DB00850). These drugs are common to both the Phenothiazines and Antipsychotic Agents groups, suggesting that they interact with every drug in the Antipsychotic Agents category. Furthermore, interactions with drugs in the dopamine group yield similar interaction results[9]. That means all the dopamine class of drugs are interacting with the subset of drugs. This is consistent with the finding from the large cluster of drug interactions in Fig. 2.

These interactions increase the risk of adverse effects such as QTc prolongation, central nervous system (CNS) depression, orthostatic hypotension, hypotension, antihypertensive activities, and reduced gastrointestinal motility. These effects[88] are commonly observed when medications with similar pharmacologic targets or mechanisms are used together, highlighting the importance of cautious prescribing practices and close monitoring of patients. Furthermore, these interactions can lead to changes in serum concentration, impacting the therapeutic efficacy, especially in antipsychotic activities. This is primarily due to the drugs sharing common metabolic pathways[89] that utilize common enzymes, leading to competition, whereby one drug may inhibit/induce enzymes that metabolize another one. Understanding these metabolic interactions is crucial for optimizing treatment outcomes and minimizing the risk of adverse effects.

New DDIs identified by PS3N

In our study, we aimed to identify novel drug-drug interactions (DDIs) using the NDD DS1 dataset[75]. This dataset comprises a substantial number of drugs, specifically 414. We trained our PS3N model using the DS1 dataset and subsequently employed it to predict potential DDIs in a similarity network constructed from the DrugBank and Protein Data Bank datasets. The similarity network was generated using protein sequence and protein structure information. It is important to note that the DrugBank dataset used for testing contains 904 drugs, providing a significantly larger network for evaluation compared to the training set.

Our predictions resulted in a total of 32,548 potential new DDIs. To refine these results, we applied a filtering step in which we eliminated interactions with predicted values below a threshold of 0.80. This filtering reduced the number of predicted DDIs to 26,359.

To better interpret the filtered predictions, we categorized the DDIs into score ranges based on their confidence levels. As shown in Table 8, a significant majority of the predicted interactions (82.45%) fall within the high-confidence range of 0.95 to 1.00. Furthermore, the data highlights an even higher concentration of DDIs in the 0.99 to 1.00 range, accounting for 62.91% of total interactions. Of these, 40.01% have scores in the 0.999 to 1.00 range, and 29.36% with a score of 1.00.

Figure 4 visualizes the distribution of these newly identified interactions. It is important to clarify that the model was trained exclusively on the DS1 dataset but was used to make predictions on the DrugBank similarity network. The filtering and ranking process described was applied to the predictions derived from this new similarity network.

Ranking New DDIs based on Similarity values

Figure 4 basically shows filtered and prioritized DDIs we discovered during the prediction process. We employed a rigorous ranking strategy that takes into account both existing similarity values and predicted values from the PS3N model.

Predicted score range	Number of interactions	Percentage of total interactions (%)
(0.8, 0.85]	966	3.67
(0.85, 0.9]	1354	5.14
(0.9, 0.95]	2305	8.75
(0.95, 1.0]	21729	82.45
(0.99, 1.0]	16582	62.91
(0.999, 1.0]	10547	40.01
[1.00 - 1.00]	7740	29.36

Table 8. Distribution of predicted scores for drug interactions using PS3N model in drugBank dataset.

- *Protien sequence similarity vs predicted similarity:* Our first ranking criterion involves a careful comparison between the prediction value of PS3N and the protein similarity value from the sequence similarity network. The underlying assumption here is that drugs with interactions are more likely to exhibit higher levels of protein interaction. We set a conservative threshold value of 0.8 to guide our filtering process. Interactions with a predicted similarity score below this threshold were not considered for further evaluation.
- *Protien structure similarity vs predicted similarity:* For the second level of ranking, we extended our comparison of the prediction value from PS3N with the similarity value from the protein structure similarity network. Similar to the first ranking, interactions having a similarity value below the similarity threshold of 0.8 were omitted from further consideration.

After carefully going through these two levels of ranking, we found a set of predicted new drug interactions, shown in Fig. 4. Our ranking and filtering approach **identified 297 completely new interactions** that represent the highest ranked candidates for potential drug interactions based on our proposed ranking criteria as showed in the Fig. 3. Selecting new predicted interactions validates the performance of the PS3N model, reinforcing the reliability and significance of these newly discovered DDIs.

Significance and clinical relevance of newly found DDI's

Figure 4 shows novel interactions identified by our PS3N approach. these have never been reported in the literature (to our knowledge). Each interaction shows the edge weights which are the prediction probability of the interactions between drugs. We can also see that here some of the drugs have a larger number of interactions which form subtree structures for them. For example, the drugs DAI DB00548, DB00350, DB00552, DB01215, DB00884, DB01586, and DB01241 formed subtrees of drug interactions.

Taking DB00548 as an example, this drug is primarily used topically in a cream formulation for the treatment of mild to moderate acne[9]. This drug interacts with 20 other drugs with different pharmacological profiles and therapeutic uses that never been found in any dataset. They include anticancer agents, antivirals, cardiovascular medications, and treatments for conditions like osteoporosis, glaucoma, and inflammation[9]. While there may not be direct pharmacological connections between DB00548 and the 20 drugs, their use in clinical practice, especially in patients with concurrent dermatological and systemic conditions, necessitates careful consideration of potential interactions, both pharmacodynamic and pharmacokinetic, to ensure safe and effective treatment outcomes.

Further, in Fig. 4, we found some interesting brand new drug interactions discovered by our model which are not included in any standard sources of drug interaction data. For example, interactions between Acetazolamide (DB00863) and Droperidol (DB00433)/Clozapine (DB00363), Brinzolamide (DB00857) and Thiothixene (DB01621), Dexmedetomidine (DB00695) and Compazine (DB00494)/Ropinirole (DB00268),

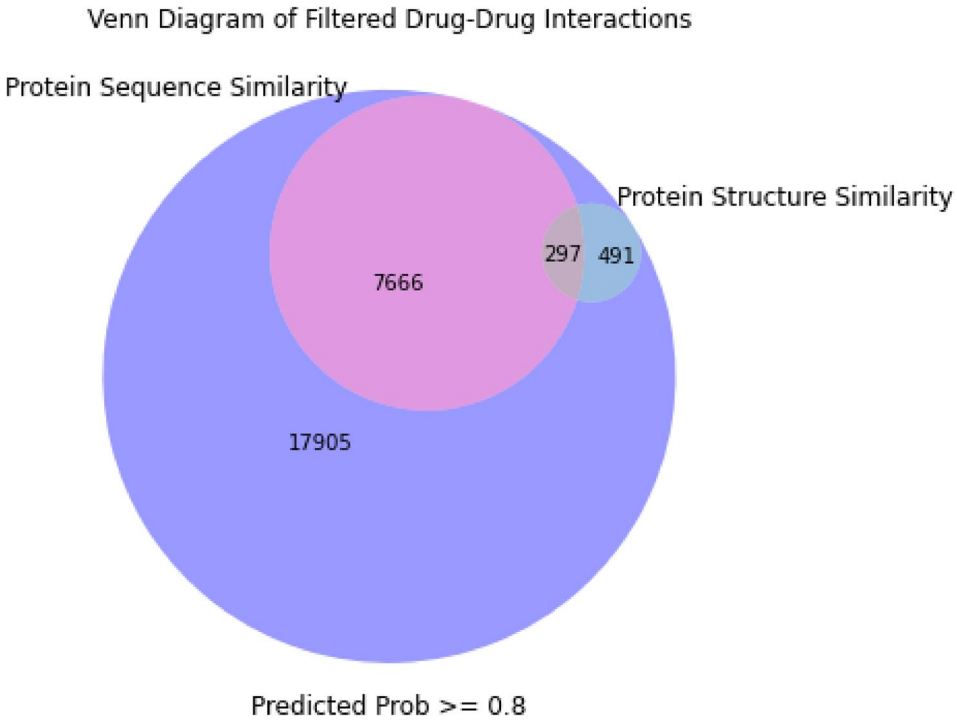

Fig. 3. Venn diagram illustrating the overlap among drug-drug interaction pairs based on three filtering criteria: (1) predicted interaction probability (Predicted Prob $\geq$ 0.8), (2) protein sequence similarity (Protein Similarity $\geq$ 0.8), and (3) protein structure similarity (PDB Similarity $\geq$ 0.8). The intersections highlight drug pairs that meet multiple filtering criteria, emphasizing the consistency across prediction and biological similarity metrics.

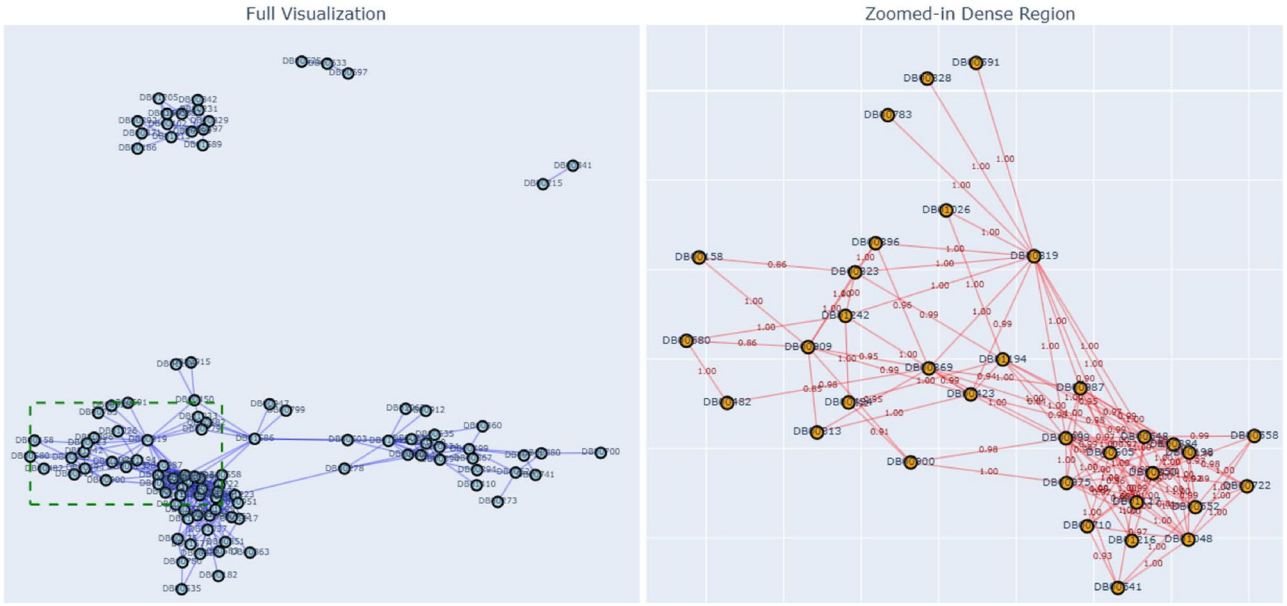

Fig. 4. Detected new DDI's from using the PS3N model using the DrugBank Dataset. The interaction network represents the DDI's which has a predicted value between 0.8 to 1.00 from the PS3N model and also passed the 2-step filtering process (297 DDI's, 904 Drugs).

Celecoxib (DB00482) and Clozapine (DB00363), Entacapone (DB00494) and Clozapine (DB00363), and Clozapine (DB00363) and Ropinirole (DB00268)/Actigall (DB01592)/Varenicline (DB01273) are noteworthy. While these medications belong to different categories, they are often used together[90] in clinical practice. For instance, Clozapine is a highly effective antipsychotic used for treatment-resistant schizophrenia, and concurrent prescriptions of Ropinirole and Varenicline are common in patients with schizophrenia and Restless Leg Syndrome[91] who are also trying to quit smoking[92]. These interactions underscore the complexity of medication management in patients with multiple comorbidities and highlight the importance of comprehensive medication reviews and interdisciplinary communication in clinical decision-making.

Understanding the clinical relevance of drug interactions is crucial for ensuring safe and effective pharmacotherapy in patients with psychiatric disorders. These findings contribute to the broader understanding of drug interactions and provide valuable insights into optimizing medication regimens for improved patient outcomes.

Discussion

The main objective of this work is to develop a new computational model for DDI prediction utilizing genetic information about drug-protein targets. To our knowledge this is the first attempt to utilize information about the protien sequences and protein structures of drug protein targets to analyze potential drug-drug interaction. Our work presents a promising approach to tackling drug-drug interaction (DDI) prediction challenges. We explored multiple methods for constructing the feature space to identify interactions between drug pairs. Specifically, we characterized drugs using protein structure and protein sequence information. To build the labeled feature space, we leveraged interaction data available from DrugBank. Combining structural and sequence-based information, we obtained a total of 904 drugs. Unlike previous methodologies, we considered only protein sequence and structure similarity networks for the first time to predict drug interactions. In addition, our similarity network computation technique allows extracting important protein features in terms of different distance measures.

Our proposed DDI framework represents a novel approach, unlike many previous studies that primarily relied on patient information and considered multiple patient circumstances. We created a rich similarity network dataset that potentially benefits various studies related to drug interactions in different clinical trials.

A limitation of our approach is that due to the different sources of data required, getting all types of information for the same drug is non trivial. Moreover, the datasets have significantly more unknown interactions than known interactions. Thus, this creates a data imbalance problem, especially if we do not consider appropriate unknown thresholds. As we mainly focused on Drugbank and Protein Data Bank (PDB), finding the commonality between the different datasets like side effects, protein protein interaction, indication data etc. also requires a lot of work. Another potential drawback is that our filtering process relied heavily on protein sequence and structure data to construct similarity networks, which might bias the ranking criteria towards interactions with high protein similarity. This approach could overlook alternative mechanisms of interaction. To mitigate this, integrating orthogonal data sources as chemical structures of drugs and phenotypic effects could provide a more comprehensive evaluation and reduce reliance on protein-centric similarity measures. Lastly, the time and space complexity for feature space generation is significant and will need to be addressed in the future.

A critical limitation arises in the context of inductive split evaluations, particularly for cold-start scenarios where drugs in the test set are unseen during training. While our model demonstrates competitive accuracy in inductive cold-split settings (CS1), its AUROC performance lags behind methods like SSI_DDI, suggesting reduced robustness in distinguishing positive and negative interactions for novel drugs. This gap highlights the need for more generalizable feature representations, possibly through advanced embedding techniques or multi-modal data fusion. Furthermore, the computational cost of generating protein-based feature spaces is non-negligible, posing scalability challenges for large-scale applications. Addressing these bottlenecks through optimized algorithms or distributed computing will be essential for real-world deployment.

In summary, while our protein-centric approach offers a novel perspective on DDI prediction, its current limitations underscore the importance of balancing protein similarity with other interaction mechanisms, improving robustness in inductive settings, and enhancing computational efficiency. Future efforts will focus on refining feature representation, incorporating auxiliary biological data, and developing scalable solutions to advance the model's applicability in clinical and pharmacological research.

Conclusion

In this study, we introduced a novel mechanism for detecting drug-drug interactions, structured into three key stages. The first stage involves constructing similarity profiles using data from DrugBank and PDB. The second stage focuses on creating an integrated similarity network (PS3N) for drugs, which incorporates information about their protein targets, specifically the protein sequences and structures. Finally, the third stage utilizes data from the integrated network to develop a deep neural network model, enhancing the prediction of potential drug interactions. We evaluated the performance of our proposed PS3N-based deep learning framework against recent machine learning approaches. The results demonstrated that our methodology is highly competitive with current state-of-the-art techniques and, in some cases, surpasses them in performance. In our proposed methodology, we showed a new approach to dealing with the DDI prediction problem by exploiting genetic information about the drug-protein targets, particularly their protein sequence and protein structure.

Our validation process reinforced the robustness of our model in detecting new drug-drug interactions (DDIs). The carefully designed analysis provided strong support for the model's performance. The DDI network diagram, in particular, offers valuable insights into drug interactions and their potential implications in identifying adverse drug events such as side effects and adverse reactions. An interesting future work would be to enhance the scope of our PS3N model to create a more generalized prediction framework for DDIs. For instance, by incorporating additional drug-related information beyond protein sequences and structures, allowing us to gain a more comprehensive understanding of drug interactions. Another is to assess the model's performance in detecting interactions specifically related to specific disease contexts, for example cardiovascular and COVID-19 drugs. We plan to utilize our model in disease-specific contexts, providing valuable insights for clinical applications.

Furthermore, our research can extend beyond DDIs to explore other adverse drug events. By applying our approach to a wider spectrum of drug-related challenges, we can contribute to a deeper understanding of drug safety and effectiveness.

Additional information

Impact of algorithmic parameters

Table 9 shows the impact of different hyperparameters on the performance of the proposed model. From the table, Adam Optimizer with a learning rate of 0.01 produced the best overall result. SGD Optimizer for learning rate 0.05, 0.01, and 0.10 showed almost similar accuracy level as we got for Adam optimizer. In our proposed neural network model, the number of hidden layers will vary based on the number of drug-active ingredients (DAIs) on the datasets. Normally, for the protein sequence dataset, it will not be more than 4. For Protein structure or the combination of both, it will be between 3 to 5.

In our proposed approach, we used similarity network fusion (SNF) to build the final matrix for each dataset. We had different distance measures, and for each measure we have one matrix. These matrices are then combined to create the fusion matrix. In the SNF method, a distance metric was considered, we used the default Euclidean distance metric. And there were two other parameters *mu*- weighted *k*-nearest neighbors kernel to the distance matrix to calculate the similarity fusion matrix. We used different values in the range of 0.3 to 0.5 for *mu* and 10 to 20 for *k*. we found that for $mu = 0.3$ and $k = 10$ works better for the performance of the model.

Optimizer	Learning rate	Accuracy
Adam Optimizer	0.05	0.7200
Adam Optimizer	0.10	0.7213
Adam Optimizer	**0.01**	**0.9710**
SGD	0.05	0.9646
SGD	0.10	0.9644
SGD	0.01	0.9637
RMSProp	0.01	0.7210

Table 9. Results of PS3N with variation on algorithmic parameters.

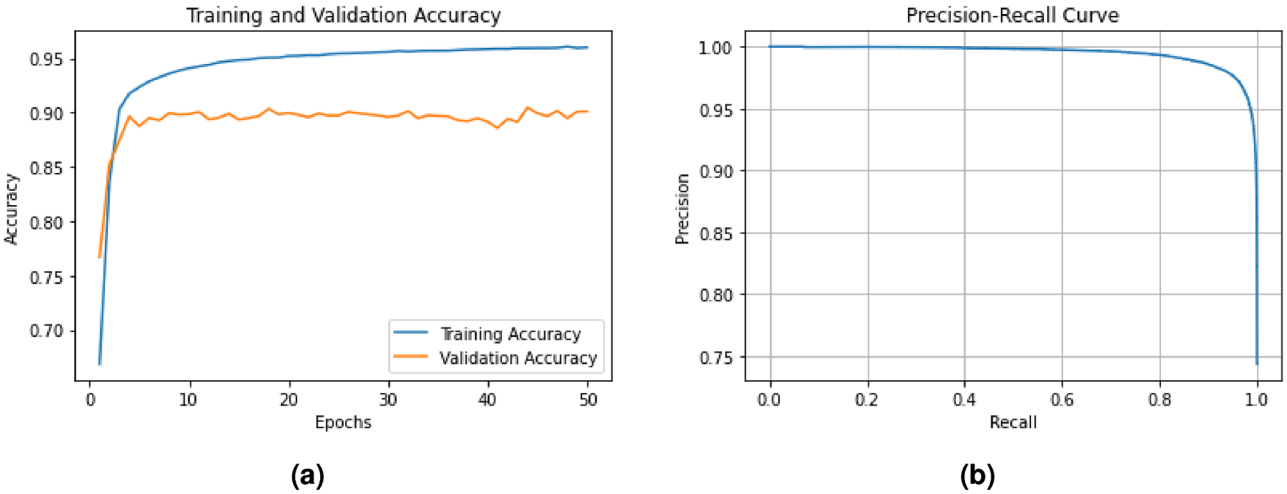

Figure 5. Model performance curves. **a** Training and validation accuracy curve. **b** Precision-Recall.

In the deep neural network model, we used 3 to 5 hidden layers based on the number of drug active ingredients present in the datasets. However, in our proposed PS3N model we used the normal distribution for the layer initialization when we used ReLU as the activation function in the layer. We used the normal because we used ReLu as activation. In the final layer, we used glorot normal distribution for the layer initialization because the Sigmoid function is used as the activation function. Though different initialization has very little impact on the performance, we select this as a general setup for our experiments.

Then we used the dropout value with a range of 0.2 to 0.5 for each of the layers. 0.3 showed the best outcome in the model performance. Then in the model optimization, we tested different optimization techniques, namely, stochastic gradient descent (SGD), Adam optimizer, RMSProp, etc in which Adam and SGD have close performance but in RMSProp, the algorithm performance decreases significantly.

In gradient calculation, the learning rate is very important. In our experimental setup, we choose different learning rates in a range from 0.001 to 0.1. We also use weight decay for the gradient update in a range from $1e^{-}4$ to $1e^{-}8$. However, we found the best value of weight decay for the model to be $1e^{-}6$. After a different selection of learning rates, we found that 0.01 gives us the best result for the model.

Model performance evaluation

Figure 5a shows the behavior of the proposed model during training and validation. We observe that throughout the training process, both training and validation accuracy maintain a relatively stable gap (of about 5%). This behavior indicates that the model effectively generalizes from the training data to previously unseen validation data. Typically, a substantial gap between the training accuracy and validation accuracy would suggest overfitting, where the model is trained well on training data but struggles to generalize. In our model, the minimal difference between the two curves demonstrates that our model's robustness. This observation provides us with confidence in the model's predictability on unseen data in practical scenarios.

In, Fig. 5b, we can further evaluate our model performance through the precision-recall curve. Our model shows a high AUC score which signifies that the model effectively identifies positive DDI while minimizing false positives which is very crucial in DDI prediction.

Data availability

Data is provided within the manuscript or supplementary information files

Received: 18 January 2025; Accepted: 1 August 2025
Published online: 24 October 2025

References

1. Edwards, I. R. & Aronson, J. K. Adverse drug reactions: Definitions, diagnosis, and management. *The Lancet* **356**, 1255–1259 (2000).
2. Pirmohamed, M. et al. Adverse drug reactions as cause of admission to hospital: Prospective analysis of 18 820 patients. *BMJ* **329**, 15–19 (2004).
3. Gottlieb, A., Stein, G. Y., Oron, Y., Ruppin, E. & Sharan, R. Indi: A computational framework for inferring drug interactions and their associated recommendations. *Mol. Syst. Biol.* **8**, 592 (2012).
4. Huang, J. et al. Systematic prediction of pharmacodynamic drug-drug interactions through protein-protein-interaction network. *PLoS Comput. Biol.* **9**, e1002998 (2013).
5. Vilar, S., Uriarte, E., Santana, L., Tatonetti, N. P. & Friedman, C. Detection of drug-drug interactions by modeling interaction profile fingerprints. *PloS One* **8**, e58321 (2013).
6. Park, K., Kim, D., Ha, S. & Lee, D. Predicting pharmacodynamic drug-drug interactions through signaling propagation interference on protein-protein interaction networks. *PloS One* **10**, e0140816 (2015).

7. Ryu, J. Y., Kim, H. U. & Lee, S. Y. Deep learning improves prediction of drug-drug and drug-food interactions. *Proc. Natl Acad Sci.* **115**, E4304–E4311 (2018).
8. Law, V. et al. Shedding new light on drug metabolism. Drugbank 4.0.. *Nucleic Acids Res.* **42**, D1091–D1097 (2014).
9. Wishart, D. S. et al. Drugbank 5.0: A major update to the drugbank database for 2018. *Nucleic Acids Res.* **46**, 1074–1082 (2018).
10. Knox, C. et al. Drugbank 3.0: A comprehensive resource for 'omics' research on drugs. *Nucleic Acids Res* **39**, D1035–D1041 (2010).
11. Kanehisa, M., Goto, S., Furumichi, M., Tanabe, M. & Hirakawa, M. Kegg for representation and analysis of molecular networks involving diseases and drugs. *Nucleic Acids Res.* **38**, D355–D360 (2010).
12. Kuhn, M., Campillos, M., Letunic, I., Jensen, L. J. & Bork, P. A side effect resource to capture phenotypic effects of drugs. *Mol. Syst. Biol.* **6**, 343 (2010).
13. Li, Q., Cheng, T., Wang, Y. & Bryant, S. H. PubChem as a public resource for drug discovery. *Drug Discovery Today* **15**, 1052–1057 (2010).
14. Wang, Y. et al. PubChem: A public information system for analyzing bioactivities of small molecules. *Nucleic Acids Res.* **37**, W623–W633 (2009).
15. Wishart, D. S. et al. DrugBank: A comprehensive resource for in silico drug discovery and exploration. *Nucleic Acids Res.* **34**, D668–D672 (2006).
16. Wishart, D. S. et al. DrugBank: A knowledgebase for drugs, drug actions and drug targets. *Nucleic Acids Res.* **36**, D901–D906 (2008).
17. Deng, Y. et al. A multimodal deep learning framework for predicting drug-drug interaction events. *Bioinformatics* **36**, 4316–4322 (2020).
18. Tanvir, F., Saifuddin, K. M., Islam, M. I. K. & Akbas, E. Ddi prediction with heterogeneous information network-meta-path based approach. *IEEE/ACM Trans. Comput. Biol. Bioinf.* **21**(5), 1168–1179 (2024).
19. Qian, S., Liang, S. & Yu, H. Leveraging genetic interactions for adverse drug-drug interaction prediction. *PLoS Computational Biol.* **15**, e1007068 (2019).
20. Lin, S. et al. MDF-SA-DDI: Predicting drug-drug interaction events based on multi-source drug fusion, multi-source feature fusion and transformer self-attention mechanism. *Brief. Bioinform.* **23**, bbab421 (2022).
21. Nyamabo, A. K., Yu, H. & Shi, J.-Y. SSI-DDI: Substructure-substructure interactions for drug-drug interaction prediction. *Brief. Bioinf.* **22**, bbab133 (2021).
22. Xiong, Z. et al. Multi-relational contrastive learning graph neural network for drug-drug interaction event prediction. *In Proceedings of the AAAI Conference on Artificial Intelligence* **37**, 5339–5347 (2023).
23. Zhang, W., Chen, Y., Li, D. & Yue, X. Manifold regularized matrix factorization for drug-drug interaction prediction. *J. Biomed. Inf.* **88**, 90–97 (2018).
24. Shi, J.-Y. et al. TMFUF: A triple matrix factorization-based unified framework for predicting comprehensive drug-drug interactions of new drugs. *BMC Bioinf.* **19**, 27–37 (2018).
25. Deepika, S. & Geetha, T. A meta-learning framework using representation learning to predict drug-drug interaction. *J. Biomed. Inf.* **84**, 136–147 (2018).
26. Han, C.-D., Wang, C.-C., Huang, L. & Chen, X. MCFF-MTDDI: Multi-channel feature fusion for multi-typed drug-drug interaction prediction. *Brief. Bioinf.* **24**, bbad215 (2023).
27. Lazarou, J., Pomeranz, B. H. & Corey, P. N. Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies. *Jama* **279**, 1200–1205 (1998).
28. Prueksaritanont, T. et al. Drug-drug interaction studies: Regulatory guidance and an industry perspective. *The AAPS J.* **15**, 629–645 (2013).
29. Kusuhara, H. How far should we go? Perspective of drug-drug interaction studies in drug development. *Drug Metabol. Pharmacokinet.* **29**, 227–228 (2014).
30. Zhang, P. et al. Translational biomedical informatics and pharmacometrics approaches in the drug interactions research. *CPT: Pharmacomet. Syst. Pharmacol.* **7**, 90–102 (2018).
31. Drug interactions: What you should know. https://www.fda.gov/drugs/resources-you-drugs/drug-interactions-what-you-should-know (2013). Accessed: 2023-10-01.
32. Lynch, S. S. Drug interactions. https://www.merckmanuals.com/professional/clinical-pharmacology/factors-affecting-response-to-drugs/drug-interactions (2019). Accessed: 2023-10-01.
33. Drug interactions. https://www.fda.gov/drugs/questions-and-answers-fdas-adverse-event-reporting-system-faers/reports-received-and-reports-entered-faers-year (2015). Accessed: 2023-10-01.
34. Rohani, N., Eslahchi, C. & Katanforoush, A. Iscmf: Integrated similarity-constrained matrix factorization for drug-drug interaction prediction. *Netw. Model. Anal. Health Inf. Bioinf.* **9**, 1–8 (2020).
35. Gan, Y., Liu, W., Xu, G., Yan, C. & Zou, G. Dmfddi: Deep multimodal fusion for drug-drug interaction prediction. *Brief. Bioinf.* **24**, bbad397 (2023).
36. Chen, Y. et al. Muffin: Multi-scale feature fusion for drug-drug interaction prediction. *Bioinformatics* **37**, 2651–2658 (2021).
37. Wang, Y.-C., Zhang, C.-H., Deng, N.-Y. & Wang, Y. Kernel-based data fusion improves the drug-protein interaction prediction. *Comput. Biol. Chem.* **35**, 353–362 (2011).
38. Shtar, G., Solomon, A., Mazuz, E., Rokach, L. & Shapira, B. A simplified similarity-based approach for drug-drug interaction prediction. *Plos one* **18**, e0293629 (2023).
39. Yildirim, M. A., Goh, K. I., Cusick, M. E., Barabási, A. L. & Vidal, M. Drug-target network. *Nature Biotechnol.* **25**, 1119–1126 (2007).
40. Gao, H. et al. High-throughput screening using patient-derived tumor xenografts to predict clinical trial drug response. *Nature Med.* **21**, 1318–1325 (2015).
41. Fang, H.-B., Chen, X., Pei, X.-Y., Grant, S. & Tan, M. Experimental design and statistical analysis for three-drug combination studies. *Stat. Methods Med. Res.* **26**, 1261–1280 (2017).
42. Cheng, F. & Zhao, Z. Machine learning-based prediction of drug-drug interactions by integrating drug phenotypic, therapeutic, chemical, and genomic properties. *J. Am. Med. Inf. Assoc.* **21**, e278–e286 (2014).
43. Zhang, P., Wang, F., Hu, J. & Sorrentino, R. Label propagation prediction of drug-drug interactions based on clinical side effects. *Sci. Rep.* **5**, 12339 (2015).
44. Lee, G., Park, C. & Ahn, J. Novel deep learning model for more accurate prediction of drug-drug interaction effects. *BMC Bioinf.* **20**, 1–8 (2019).
45. Kastrin, A., Ferk, P. & Leskosek, B. Predicting potential drug-drug interactions on topological and semantic similarity features using statistical learning. *PloS One* **13**, e0196865 (2018).
46. Takeda, T., Hao, M., Cheng, T., Bryant, S. H. & Wang, Y. Predicting drug-drug interactions through drug structural similarities and interaction networks incorporating pharmacokinetics and pharmacodynamics knowledge. *J. Cheminf.* **9**, 1–9 (2017).
47. Zhang, Y. et al. Drug-drug interaction extraction via hierarchical rnns on sequence and shortest dependency paths. *Bioinformatics* **34**, 828–835 (2018).
48. Yu, H. et al. Predicting and understanding comprehensive drug-drug interactions via semi-nonnegative matrix factorization. *BMC Syst. Biol.* **12**, 101–110 (2018).
49. Lü, L., Pan, L., Zhou, T., Zhang, Y.-C. & Stanley, H. E. Toward link predictability of complex networks. *Proc. Natl Acad. Sci.* **112**, 2325–2330 (2015).

50. Abbasi, A. et al. Social media analytics for smart health. *IEEE Intell. Syst.* **29**, 60–80 (2014).
51. Adjeroh, D. et al. Signal fusion for social media analysis of adverse drug events. *IEEE Intell. Syst.* **29**, 74–80 (2014).
52. Ahmad, F., Abbasi, A., Kitchens, B., Adjeroh, D. & Zeng, D. Deep learning for adverse event detection from web search. *IEEE Trans. Knowl. Data Eng.* **34**, 2681–2695 (2020).
53. Harada, S. et al. Dual graph convolutional neural network for predicting chemical networks. *BMC Bioinf.* **21**, 1–13 (2020).
54. Wang, Y., Min, Y., Chen, X. & Wu, J. Multi-view graph contrastive representation learning for drug-drug interaction prediction. *In Proceedings of the web conference* **2021**, 2921–2933 (2021).
55. Lin, X., Quan, Z., Wang, Z.-J., Ma, T. & Zeng, X. Kgnn: Knowledge graph neural network for drug-drug interaction prediction. *In IJCAI* **380**, 2739–2745 (2020).
56. Deng, Y. et al. A multimodal deep learning framework for predicting drug-drug interaction events. *Bioinformatics* **36**, 4316–4322 (2020).
57. Shang, Y., Gao, L., Zou, Q. & Yu, L. Prediction of drug-target interactions based on multi-layer network representation learning. *Neurocomputing* **434**, 80–89 (2021).
58. An, Q. & Yu, L. A heterogeneous network embedding framework for predicting similarity-based drug-target interactions. *Brief. Bioinf.* **22**, bbab275 (2021).
59. Yan, X.-Y., Yin, P.-W., Wu, X.-M. & Han, J.-X. Prediction of the drug-drug interaction types with the unified embedding features from drug similarity networks. *Front. Pharmacol.* **12**, 794205 (2021).
60. Lv, Q., Zhou, J., Yang, Z., He, H. & Chen, C.Y.-C. 3d graph neural network with few-shot learning for predicting drug-drug interactions in scaffold-based cold start scenario. *Neural Netw.* **165**, 94–105 (2023).
61. Lv, Q., Chen, G., Yang, Z., Zhong, W. & Chen, C.Y.-C. Meta learning with graph attention networks for low-data drug discovery. *IEEE Trans. Neural Netw. Learn. Syst.* **35**, 11218–11230 (2023).
62. Lv, Q., Chen, G., Yang, Z., Zhong, W. & Chen, C.Y.-C. Meta-molnet: A cross-domain benchmark for few examples drug discovery. *IEEE Trans. Neural Netw. Learn. Syst.* **36**, 4849–4863 (2024).
63. Protein data bank. The single global archive for 3d macromolecular structure data. *Nucleic Acids Res.* **47**, D520–D528 (2019).
64. Islam, S. et al. Detecting drug-drug interactions using protein sequence-structure similarity networks. In *2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 3472–3477 (IEEE, 2021).
65. Bull, S. C. & Doig, A. J. Properties of protein drug target classes. *PloS One* **10**, e0117955 (2015).
66. Clavijo, B. J. k-mer counting, part i: Introduction. https://bioinfologics.github.io/post/2018/09/17/k-mer-counting-part-i-introduction/ (2018). Accessed: 2023-10-01.
67. Yao, Y.-H. et al. Protein sequence information extraction and subcellular localization prediction with gapped k-mer method. *BMC Bioinf.* **20**, 1–8 (2019).
68. Déraspe, M., Boisvert, S., Laviolette, F., Roy, P. H. & Corbeil, J. Flexible protein database based on amino acid k-mers. *Sci. Rep.* **12**, 9101 (2022).
69. Zhang, Y., Wen, J. & Yau, S.S.-T. Phylogenetic analysis of protein sequences based on a novel k-mer natural vector method. *Genomics* **111**, 1298–1305 (2019).
70. Viscardi, L. H. et al. Foxp in tetrapoda: Intrinsically disordered regions, short linear motifs and their evolutionary significance. *Genet. Mol. Biol.* **40**, 181–190 (2017).
71. Nugent, C. M. & Adamowicz, S. J. Alignment-free classification of coi DNA barcode data with the python package alfie. *Metabarcoding Metagenomics* **4**, e55815 (2020).
72. Adjeroh, D., Bell, T. & Mukherjee, A. *The Burrows-Wheeler Transform: Data Compression, Suffix Arrays, and Pattern Matching* (Springer Science & Business Media, Berlin Heidelberg, 2008).
73. Tan, J. & Adjeroh, D. Protein family structure signature for multidomain proteins. *Int. J. Data Mining Bioinf.* **20**, 285–302 (2018).
74. Wang, B. et al. Similarity network fusion for aggregating data types on a genomic scale. *Nature Methods* **11**, 333–337 (2014).
75. Rohani, N. & Eslahchi, C. Drug-drug interaction predicting by neural network using integrated similarity. *Sci. Rep.* **9**, 13645 (2019).
76. Gomez, R. Understanding categorical cross-entropy loss, binary cross-entropy loss, softmax loss, logistic loss, focal loss and all those confusing names. https://gombru.github.io/2018/05/23/cross_entropy_loss/ (2018). Accessed: 2023-10-01.
77. Kingma, D. P. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014).
78. Zhang, W. et al. Predicting potential drug-drug interactions by integrating chemical, biological, phenotypic and network data. *BMC Bioinf.* **18**, 1–12 (2017).
79. Ren, Z.-H. et al. BioDKG-DDI: Predicting drug-drug interactions based on drug knowledge graph fusing biochemical information. *Brief. Funct. Genomics* **21**, 216–229 (2022).
80. Feng, Y.-H., Zhang, S.-W. & Shi, J.-Y. DPDDI: A deep predictor for drug-drug interactions. *BMC Bioinformatics* **21**, 419 (2020).
81. Lin, J. et al. R2-DDI: Relation-aware feature refinement for drug-drug interaction prediction. *Brief. Bioinf.* **24**, bbac576 (2023).
82. Nyamabo, A. K., Yu, H., Liu, Z. & Shi, J.-Y. Drug-drug interaction prediction with learnable size-adaptive molecular substructures. *Brief. Bioinf.* **23**, bbab441 (2022).
83. Tang, Z., Chen, G., Yang, H., Zhong, W. & Chen, C.Y.-C. DSIL-DDI: A domain-invariant substructure interaction learning for generalizable drug-drug interaction prediction. *IEEE Trans. Neural Netw. Learn. Syst.* **35**(8), 10552–10560 (2023).
84. Gao, J., Wu, Z., Al-Sabri, R., Oloulade, B. M. & Chen, J. AutoDDI: Drug-drug interaction prediction with automated graph neural network. *IEEE J. Biomed. Health Inf.* **28**(3), 1773–1784 (2024).
85. Ryu, J. Y., Kim, H. U. & Lee, S. Y. Deep learning improves prediction of drug-drug and drug-food interactions. *Proc. Natl Acad. Sci.* **115**, E4304–E4311 (2018).
86. Xu, N., Wang, P., Chen, L., Tao, J. & Zhao, J. Mr-gnn: Multi-resolution and dual graph neural network for predicting structured entity interactions. arXiv preprint arXiv:1905.09558 (2019).
87. Deac, A., Huang, Y.-H., Veličković, P., Liò, P. & Tang, J. Drug-drug adverse effect prediction with graph co-attention. arXiv preprint arXiv:1905.00534 (2019).
88. Demler, T. L. Exploring psychotropic drug interactions. *Psychiatric Times* **37** (2020).
89. Preskorn, S. H. Drug-drug interactions (ddis) in psychiatric practice, part 3: Pharmacokinetic considerations. *J. Psychiatr. Pract.* **25**, 34–40 (2019).
90. Nucifora, F. C. Jr., Mihaljevic, M., Lee, B. J. & Sawa, A. Clozapine as a model for antipsychotic development. *Neurotherapeutics* **14**, 750–761 (2017).
91. Gossard, T. R., Trotti, L. M., Videnovic, A. & St Louis, E. K. Restless legs syndrome: Contemporary diagnosis and treatment. *Neurotherapeutics.* **18**(1), 140–55 (2021).
92. Rigotti, N. A., Kruse, G. R., Livingstone-Banks, J. & Hartmann-Boyce, J. Treatment of tobacco smoking: a review. *Jama* **327**, 566–577 (2022).

Acknowledgements

This work was supported in part by grants from the US National Science Foundation (Awards #1816005, #2039915, #1920920).

Author contributions

Must include all authors, identified by initials, for example: A.A. conceived the experiment(s), A.A. and B.A. conducted the experiment(s), C.A. and D.A. analysed the results. All authors reviewed the manuscript.

Funding

This work was supported in part by the US National Science Foundation under Awards #1816005, #2039915, and #1920920.

Declarations

Competing interest

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to S.I.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.