Received: 24 November 2024 | *Revised:* 10 August 2025 | *Accepted:* 16 August 2025
DOI: 10.37917/ijeee.22.1.40 *Vol. 22 | Issue 1 | June 2026*

Iraqi Journal for Electrical and Electronic Engineering
Original Article

Open Access
IJEEE University of Basrah College of Engineering

An Attention-Based Graph Neural Network Method for Drug-Drug Interaction Prediction

Murteza Hanoon Tuama*[1,2], Amir Lakizadeh[1]
[1]Department Computer Engineering and Information Technology, University of Qom, Qom, Iran
[2]Department of Computer Techniques Engineering, Imam Al-Kadhum College (IKC), Iraq

Correspondance
*Murteza Hanoon Tuama
Department Computer Engineering and Information Technology, University of Qom, Qom, Iran
Email: murtezahanoon@iku.edu.iq

Abstract
Drug-drug interactions (DDIs) stand at the forefront of challenges in modern pharmacology, necessitating precise prediction methods to ensure patient safety. This study presents a pioneering approach that synergizes attributed heterogeneous graph embedding with deep learning to forecast DDIs and their specific classifications. Our methodology is delineated into two pivotal stages. The preliminary phase revolves around data assimilation, leading to the creation of specialized feature matrices such as Chemical Composition, Interaction Targets, Enzymatic Reactions, and Biological Pathways. These matrices culminate in a comprehensive drug network where drugs are symbolized as nodes. Upon rigorous data refinement, these matrices serve as attribute markers for each node. Capitalizing on the robustness of the attributed heterogeneous network, we amalgamate diverse drug attributes, thereby amplifying the depth of drug interaction assessments. The subsequent phase sees these drug embedding vectors undergo strategic concatenation, resulting in detailed feature vectors for drug pairings. The final step involves a dense neural network, tasked with decoding intricate drug interaction nuances. The introduction of an attention-driven embedding process further accentuates the model's capability by emphasizing pivotal interactions. The promising results, coupled with an innovative methodology, sets the stage for future explorations, potentially revolutionizing DDI predictions.

Keywords
Attention mechanism, Drug-Drug interaction, Deep learning, Graph neural network.

I. Introduction

Patients with chronic illnesses like hypertension, cancer, and cardiac issues frequently take multiple medications at the same time. This concurrent consumption can lead to DDIs, potentially diminishing the effectiveness of the treatments and possibly causing adverse drug reactions (ADRs) [1, 2]. Recognizing these potential DDIs early on is crucial to decrease negative outcomes. Nonetheless, due to the limited scale and duration of clinical trials, many such interactions might go unnoticed before the drug reaches the market [3]. Carrying out clinical tests for every possible drug combination is also prohibitively expensive. Consequently, there's a pressing need for a digital approach that can automatically detect and address unexpected DDIs.

Public databases such as DrugBank1, STITCH2, SIDER3, PubChem4, KEGG5, among others, have paved the way for various computational models designed to identify Drug-Drug Interactions (DDIs) [4, 5]. Many of these models assess the SMILES (Simplified Molecular Input Line Entry System) similarity or binding characteristics of drug pairs [6]. SMILES is a unique notation that employs ASCII symbols to represent molecular configurations in detail. A sequence of these symbols can represent a molecule's three-dimensional chemical architecture. However, within a molecule's entire structure, only certain substructures play a pivotal role in the chemical interactions between drugs, rendering the remaining structures relatively inconsequential [7]. Focusing on the full chemical structure could give undue emphasis to these non-

essential substructures, potentially weakening the accuracy of DDI predictions [8]. Given our growing capability to access relational drug information, contemporary techniques often amalgamate diverse data sources to identify drug attributes, such as side effects, target proteins, pathways, and therapeutic indications [3, 9].

Recently, there's been a surge in interest in using network-based techniques for the study of DDIs. The bulk of these graph- centered approaches focus on binary relationships between drugs. Typically, in these methods, a straightforward graph is used where each node represents a drug and edges indicate known interactions between these drugs. Nonetheless, some strategies delve deeper, factoring in relationships between drugs and other biological factors, leading to the construction of heterogeneous networks. From these networks, varying topological data is harnessed to anticipate unidentified connections or interactions between drugs. Over recent years, deep learning has revolutionized many sectors, with the pharmaceutical domain being no exception. This advanced subset of machine learning has brought about a paradigm shift in drug discovery and DDI prediction. Traditional methods often relied on structured data and specific features for drug interactions, but deep learning has the ability to automatically extract patterns and features from vast amounts of unstructured data. Neural networks, particularly convolutional and recurrent networks, have been at the forefront of this transformation. By leveraging vast datasets and computational power, these networks analyze molecular structures, physiological reactions, and other drug properties, offering a more holistic view of potential interactions. This leap in technology has not only expedited the process of drug discovery but has also improved the accuracy and reliability of DDI predictions, providing a safer and more efficient pathway to drug development. However, as with any technology, deep learning models for DDI have had their challenges, leading to the exploration of even more advanced techniques like graph neural networks (GNN) to further refine the process.

With the emergence and enhancement of graph neural networks (GNN), several models using GNN for DDI prediction have been put forward [5, 10]. Some models manually generate heterogeneous graphs drawing from various sources, while others

form biomedical knowledge graphs by distilling triplets from raw datasets like DrugBank [11, 12]. In these intricate networks, different components. The connections or relationships among these components are portrayed as edges. Although integrating diverse drug-related details from multiple sources can provide a deeper understanding of DDIs and has shown promising results, merging data from various origins poses challenges. Deciphering which piece of information is most critical for DDI prediction requires profound knowledge of biomedical entities. This becomes even more complex for drugs in the nascent stages of development. Additionally, having comprehensive data for every drug is often unrealistic, especially for newly introduced ones. Hence, these models might falter when any piece of data goes missing [13].

In this study, we introduce an innovative approach for forecasting DDIs and their respective categories utilizing a blend of attributed heterogeneous graph embedding combined with deep learning techniques. Our methodology is bifurcated into two distinct phases. The initial phase emphasizes the accumulation of data, subsequently leading to the formulation of four distinct feature matrices. Additionally, a matrix specific to drug-drug interactions is devised. This matrix subsequently aids in the construction of a multifaceted drug network, where each drug is represented as a node. Upon meticulous preprocessing, the feature matrices are integrated into this network, serving as attribute descriptors for each node. Leveraging the capabilities of the attributed heterogeneous network representation, we seamlessly fuse diverse drug attributes to enhance the granularity of drug interaction classifications, thereby facilitating the generation of insightful drug embedding vectors.

Transitioning to the second phase, these embedding vectors corresponding to individual drugs are readied. By implementing a strategic concatenation approach, we generate comprehensive feature vectors for drug pairs. The culmination of this process involves deploying a densely connected neural network, which harnesses these vectors to discern the nuances of drug interaction categories. Augmenting our methodology, we integrate an attention-centric embedding procedure within the established graph. This technique assigns proportional weights to each node (representing drugs) predicated on its significance, thereby enabling the model to prioritize crucial interactions throughout the prediction trajectory.

II. Related Works

Over the years, various methodologies have been put forward to tackle the DDI prediction challenge. These approaches can generally be segmented into three primary categories: similarity-based, classification-based, and network-based methods. Historically, a common hypothesis in similarity-based methods is that drugs with analogous characteristics are likely to have comparable interaction profiles. As such, many techniques strive to pinpoint resemblances between drugs. Various metrics have been devised to measure these resemblances, such as pharmacological, topological, or semantic similarities, with these metrics often rooted in statistical learning, as shown in Table I.

For instance, Vilar et al. [14] ventured into predicting DDIs by tapping into molecular similarities. They encapsulated each drug using a molecular fingerprint, which is essentially a binary vector, signaling the existence (or non-

TABLE I.
TYPE SIZE FOR PAPERS

Approach	Reference	Main Idea and Techniques
Similarity-Based	Vilar et al. [14]	Molecular similarities with binary vector fingerprints.
	INDI [15]	Amalgamates seven measures of drug-drug similarity from varied data sources.
	Study [16]	Uses four distinct biological information types for drug similarity.
Classification-Based	Davazdahemami and Delen [17]	Graph incorporating drug-protein and drug-side effect interactions, with classification strategy.
	Luo et al. [18]	Instantaneous DDI predictions based on molecular configuration, with logistic regression model.
	Ibrahim et al. [19]	Variety of similarity features for ML classifiers.
Network-Based	Study [20]	Drug networks based on confirmed DDIs.
	Study [21]	Molecular graph construction from SMILES notation.
	Decagon [5]	Knowledge graph with relational GCNN for multi-relational, multimodal networks.

existence) of specific molecular attributes. In another intriguing study, [15] introduced INDI, a system which amalgamates seven varied measures of drug-drug similarity gleaned from a spectrum of data including drug side-effects, molecular fingerprints, therapeutic effects, and more. A noteworthy contribution also comes from [16], where the researchers factor in a quartet of distinct biological information types to gauge the similarity between drug duos.

Numerous models leverage features drawn from diverse biological entities and drug interaction data, subsequently applying varied machine learning (ML) techniques for DDI training [8, 17–19].

For example, Davazdahemami and Delen [17] formulated a graph that encapsulates both drug-protein and drug-side effect interactions. Building on this, they harnessed a classification strategy on the amassed feature set, generating a series of similarity and centrality metrics rooted in the network. These metrics were then channeled into four distinct ML models for evaluation. In a different approach, Luo et al. [18] introduced a DDI prediction platform capable of delivering instantaneous predictions, relying solely on the molecular configuration of drugs. This system evaluates a drug's chemical makeup against 611 human proteins, resulting in a 611-dimensional vector that represents how the drug docks with each protein. Features for drug pairs are then formed by linking these docking vectors. Using these attributes, they then developed a logistic regression model dedicated to DDI predictions.

Taking another tack, Ibrahim et al. [19] commenced by extracting a variety of similarity features, using logistic regression to hone in on the most optimal feature. This handpicked feature was then integrated into six disparate ML classifiers with the objective of DDI prediction. However, a recurring challenge in DDI prediction pertains to the scarcity of negative samples. A novel solution to circumvent this predicament is presented in [8], termed DDI-PULearn. This method initially crafts negative samples by leveraging techniques like one-class SVM and kNN. Following this, both the positive and the synthetically created negative samples are employed in the process of DDI prediction.

In the past ten years, network-based models have gained significant traction in addressing drug-related challenges. Some researchers have designed drug networks based on confirmed DDIs, where individual drugs are represented as nodes and any interaction between drugs is signified by a connecting link [20]. Furthermore, there's a rising trend in utilizing heterogeneous information networks that incorporate diverse biomedical entities, such as proteins, side effects, and pathways, to tackle analogous issues [9].

An alternative approach is presented by [21], where a unique molecular graph is constructed for each drug derived from its SMILES notation. In addition, prevalent network-based models frequently harness drug embeddings and aim to directly discern latent node embeddings via multiple embedding techniques. However, one of their inherent limitations is the constrained capability to capture specific neighborhood information about any entity within Knowledge Graphs (KG). In recent times, Graph Neural Networks (GNNs) have emerged as powerful tools, exhibiting noteworthy results across various domains. These include areas such as drug discovery [22],

drug misuse identification [23], and drug-drug interaction prediction [21, 24, 25], among others.

For instance, Decagon [5] pioneered the development of a knowledge graph. Building on this, they conceptualized a relational GCNN, tailored specifically for predicting links across multi-relational, multimodal networks. In an innovative twist, they employed a unique graph auto-encoder method, facilitating an end-to-end trainable model adept at predicting links on a multimodal graph.

Taking another innovative route, CASTER [7] ventured into creating a dictionary learning architecture, which was geared towards predicting DDIs, with a primary focus on drug chemical structures. By capitalizing on the text representation of a drug's molecular configuration, namely its SMILES ' [26] strings, CASTER was able to surpass numerous advanced deep learning models, including the likes of DeepDDI [27] and molVAE [28]. The secret sauce of CASTER is its sequential pattern mining approach, which, in the training phase, identifies similar substrings within the provided SMILES strings. These identified patterns are then transposed into embeddings via an encoder module. Subsequently, these feature embeddings are converted into linear coefficients, which are processed through a decoder and a predictor module to generate DDI predictions.

Another noteworthy contribution comes from [25], where they devise a drug network, connecting two drugs if they bear similar chemical substructures. Following this, various GNN models are applied to this network to obtain drug representations. These representations of drug pairs are then directed to a machine learning classifier to ascertain potential interactions.

III. Materials and Methods

In this study, we introduce a dual-stage framework that amalgamates various drug attributes to anticipate DDI-specific events, leveraging the potency of attributed heterogeneous network representations in conjunction with a series of deep neural networks. In the initial phase, our framework employs a GNN model to derive drug embeddings from the attributed heterogeneous networks. Following this, the aggregated feature vectors are subjected to a series of deep neural networks to pinpoint DDI events.

Crucially, our approach incorporates an attention-driven embedding mechanism within the graph's architecture. This mechanism systematically allocates weights to individual nodes (drugs), contingent on their contextual significance, ensuring the model prioritizes pivotal interactions throughout the predictive analysis. The data preparation phase is geared towards crafting these attributed heterogeneous networks, after which representation vectors for the drugs are generated via the GNN model, serving as the foundation for learning in attributed heterogeneous networks. Subsequently, these embedding vectors are concatenated, forming the basis for training the model. The endgame is to adeptly predict both the presence of DDIs and their respective categories.

A. Data Collection

In the context of this investigation, the primary dataset was sourced from the research conducted by Deng et al. [3]. In their exploration, the team meticulously gathered and processed requisite data from esteemed databases, prominently DrugBank42 and KEGG43. This dataset encompasses four distinct drug property matrices: Chemical composition, Interaction Targets, Enzymatic reactions, and Biological Pathways. Our team secured the Pathway matrix from both the DrugBank and KEGG repositories. Conversely, the remaining matrices were exclusively extracted from DrugBank. Within these matrices, the columns symbolize the drugs, while the rows depict specific attributes related to the drugs (e.g., the variety of enzymes). A binary representation (1 or 0) within these matrices signifies the presence or absence of a distinct attribute, such as a specific enzyme linked to a drug. Furthermore, the dataset proffers an edge list detailing drug-drug relationships, encapsulating 65 varied drug-drug interaction types. These interactions in the database are articulated as a quadruple ensemble comprising drug A, drug B, the interaction mechanism, and its subsequent action. Here, the "mechanism" delineates the metabolic or therapeutic implications of the drug interactions, while "action" conveys the augmentation or diminution of these effects. For our purposes, we utilized the initial two components to structure the drug interaction edge list, and the latter pair was leveraged to define the interaction category or what we refer to as an "event". It's noteworthy that the event distribution in this dataset is skewed, indicating an imbalance in data representation.

The version 5.1.3 of DrugBank provided us with a comprehensive dataset structured into four tables:

1. The "drug" table, cataloging 572 distinct drugs alongside their attributes.
2. The "event" table, showcasing 37,264 DDIs spanning the aforementioned 572 drug varieties.
3. The "extraction" table, representing the outcomes of NLP processes, whereby each interaction is presented as a tuple containing the mechanism, action, drug A, and drug B.
4. Lastly, the "event numer" table enumerates the various DDI events and their respective occurrence frequencies.

The Table II provides a concise summary of the dataset utilized in the study. It outlines the distinct types of data and their corresponding counts, as shown in Fig. 1.

TABLE II.
SUMMARY OF THE DATASET

Type of Data	Count
Drugs	572
Drug-Drug links	37,269
Target	1,162
Enzyme	202
Chemical structure	881
Drug Pathway	957

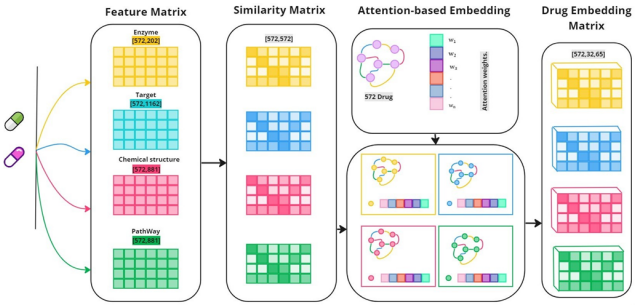

Fig. 1. Embedding matrix

The dataset contains information on 572 different drugs. Interactions between these drugs are represented through 37,269 drug-drug links. Additionally, various attributes associated with these drugs are further categorized and enumerated. This includes 1,162 different targets, 202 enzymes, 881 distinct chemical structures, and 957 unique drug pathways. The variety and depth of the dataset emphasize its comprehensive nature, offering a broad perspective on drug interactions and their associated properties.

For a comprehensive understanding of the embeddings that represent drug interactions, we harness a blend of the Jaccard similarity function and a graph-centric neural network model, augmented by an attention mechanism. Here's a more granular breakdown:

1) Jaccard Similarity Function

Given two binary feature vectors A and B, each comprised of n elements:

- M_{11} accounts for occurrences where both vectors A and B register a value of 1.
- M_{01} stands for scenarios wherein A logs a 0 and B records a 1.
- M_{10} pinpoints situations where A holds a 1 and B notes a 0.

To discern the similarity between two drug entities, let's call them d_i and, especially when pivoting around target proteins, we extract the relevant rows from the feature matrix and channel them into the Jaccard function.

$$B = \frac{M_{11}}{M_{11} + M_{01} + M_{10}} \tag{1}$$

B. Attention-Based Embedding Process

Our underlying ambition here is to implement an attention-driven embedding mechanism on the structured graph. This mechanism judiciously allots weights to the various drug nodes (or vertices) in alignment with their importance or relevance. Such a strategy ensures that our model remains acutely attuned to the most pivotal interactions, significantly enhancing the precision of subsequent predictions.

C. Constructing the Embedding

To this end, we have developed an attributed heterogeneous network which is sculpted out of the drug-drug edge list. This multifaceted network, populated with diverse edge types, lucidly presents the nature and specifics of drug interactions. Herein, nodes epitomize drugs, while the attributes tethered to these nodes mirror the unique characteristics of each drug.

In our quest to achieve an accurate drug representation, we operationalize the attributed heterogeneous network. For every iterative step, a distinct similarity matrix is annexed as a node attribute. This systematic approach culminates in the formation of four discrete networks. The act of embedding takes into account both the architectural blueprint of the network and the vectors associated with each node feature. The ensuing embedding matrix, thus formulated, encapsulates three crucial dimensions: the node count, the span of the embedding, and the diversity of edge types.

D. GNN Approach

In the vast realm of neural networks, "embedding" often refers to crafting condensed vectors for input entities, underpinned by their inherent attributes. A plethora of techniques are at our disposal to materialize such graph embeddings. Our methodology of choice Dimensionality Reduction and Deep Learning Prediction dovetails a GNN model specifically calibrated for deciphering attributed heterogeneous networks. The endgame? Extracting a distilled representation of the nodes.

Drawing inspiration from contemporary research literature, our method seeks to obtain meaningful embeddings from the said heterogeneous network. This ingenious algorithm is adept at uncovering latent attributes intrinsic to the network's structure, all while giving due weightage to node-specific attributes. Every node, represented as v_i, is bestowed with a dual embedding – a foundational base embedding and a nuanced edge embedding. These are generated by a transformative function that factors in both the node's attributes and the overarching network schema.

The modus operandi here is straightforward. The GNN ingests the heterogeneous network, complemented by attributes specific to nodes. The initial phase involves concocting training samples across different edge types via the Random Walk diffusion paradigm. Post this, the model conjures node sequences with the help of Random Walk, subsequently employing the Skip-gram technique over these sequences to glean the embeddings. This iterative process ensures the model is perpetually updated, ensuring a holistic embedding for every node across varying edge types.

Assuming we're navigating through n drugs and r edge types, and when using the Enzyme similarity matrix as a guiding node attribute, our embedding materializes as E_{in}^{er}. Additional matrices, denoted by E_{in}^{tr}, E_{in}^{pr}, and E_{in}^{sr}, are developed using n. Target, Pathway, and Chemical Structure Similarity Matrix as node characteristics, in that order. Essentially, four matrices are produced by the embedding process; they are derived from the drug-drug interactions network and the four similarity matrices associated with drug characteristics.

E. Dimensionality Reduction and Deep Learning Prediction

Upon the derivation of the embedding matrices for individual drugs, our methodology encompasses an aggregation (or concatenation) process. The primary objective here is to compress these multi-dimensional matrices into a singular, one-dimensional feature vector. This streamlined vector is then fed into a multi-layered, fully connected deep learning model, equipping it to accurately predict various DDI types.

F. Dimensionality Reduction of Embedding Matrice

s Once the network's embedding matrix has been formulated, each drug manifests as a two-dimensional matrix. Intrinsically, this matrix encapsulates the node (Drug) embedding vectors across various interaction (or edge) types. To synthesize these drug embedding vectors, we employ a concatenation process. The resultant one-dimensional vector serves as a comprehensive representation of drug i, spanning all edge types. Subsequently, the DDI list is enriched by obtaining a characteristic vector for every drug pair—achieved by multiplying the feature vectors of drugs i and. Let's elucidate further: Given that M_i is the embedding matrix of drug i and vector v in a specific edge type t is denoted as v_i,t, the one-dimensional feature vector F_i for drug i is articulated as $F_i = [v_i, 1, \ldots, v_{i,t}]$. The feature vector characterizing the drug pair is represented by $F_k = F_i \oplus F_j$, where $\oplus$ signifies the element-wise product.

G. DDI Prediction via Deep Learning

With the successful generation of four distinctive matrices of feature (or embedding) vectors, we employ a fully connected deep learning network to embark on the prediction endeavor. As portrayed in Fig. 2, the model delineated for this step encompasses four dedicated sub-networks.

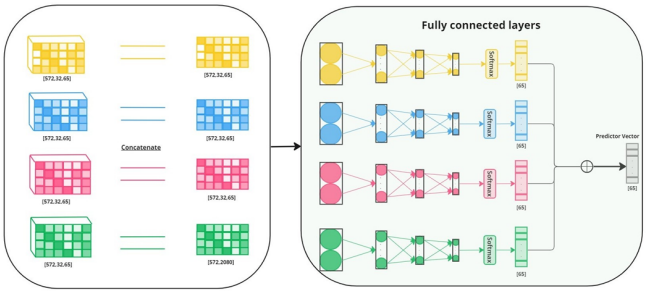

Fig. 2. Multimodal deep neural network

These are conceptualized drawing inspiration from the bottleneck neural network framework.

Each sub-network is primed to utilize one of the four matrices encompassing the drug's feature vectors as its input. The cumulative outputs of these sub-networks are then cohesively aggregated to determine the final outcome.

A plethora of hidden layers enrich the network, punctuated with batch normalization layers interspersed between them. A softmax layer is pivotal in orchestrating predictions in these sub-networks. To bolster the model's generalization capabilities while curtailing the risk of over-fitting, strategic dropout layers are integrated.

Throughout the network, the Rectified Linear Unit (ReLU) serves as the chosen activation function. The culmination of outputs from the sub-networks undergoes an averaging process, leading to the final prediction.

In terms of optimization, we hinge on the cross-entropy loss function, supplemented by the Adam optimizer, provisioned with default parameters. For a harmonious blend of preventing overfitting while expediting the training journey, the early stopping strategy is adopted. If the model doesn't register any enhancements over 10 epochs, the training is autonomously terminated

IV. Experiment Results

A. Evaluation Metrics of the Model

The results obtained from our analysis provide various evaluation metrics for the model. These metrics include Accuracy, ROC AUPR, ROC AUC, F1-score, Precision, and Recall. These metrics aid in determining the model's overall performance in predicting drug-drug interactions.

The Table III presents five sets of evaluation results.

Here are some key observations from the data:

- The accuracy of the model is consistently high, with the highest being approximately 0.9906, indicating that the model has a robust ability to correctly predict drug-drug interactions.

TABLE III.
MODEL EVALUATION METRICS

seq	Accuracy	ROC AUPR	ROC AUC	F1	Precision	Recall
0	0.900467	0.847863	0.906115	0.829244	0.756108	0.918043
1	0.942679	0.902662	0.939572	0.892447	0.855075	0.933235
2	0.944826	0.828072	0.870809	0.807671	0.855868	0.764612
3	0.985026	0.885895	0.931873	0.881629	0.892612	0.870914
4	0.990608	0.865486	0.890116	0.854288	0.941284	0.782012

- ROC AUC, which measures the overall performance of a model across all classification thresholds, is also considerably high across all iterations. This suggests that the model possesses a strong discriminative capacity.
- The F1-score, a metric that takes into account both precision and recall, displays good results, signifying a balanced model performance regarding false positives and false negatives.
- Precision, which indicates the proportion of positive identifications that were indeed correct, is also commendable, with the highest precision recorded at about 0.9413.
- Recall values are largely high as well, denoting the model's capability to identify all actual positives.

In Fig. 3 summary, the presented evaluation metrics underscore the model's efficacy in predicting drug-drug interactions with considerable accuracy, precision, and recall. The model showcases a strong ability to distinguish between interactions, making it valuable for its intended purpose.

B. Distribution of Model Evaluation Metrics

Upon visual inspection of the histogram plots for each evaluation metric, we can gather a clearer understanding of their distribution and frequency.

The accuracy of the model is consistently high, with the highest being approximately 0.9906, indicating that the model has a robust ability to correctly predict drug-drug interactions.

ROC AUC, which measures the overall performance of a model across all classification thresholds, is also considerably high across all iterations. This suggests that the model possesses a strong discriminative capacity.

The F1-score, a metric that takes into account both precision and recall, displays good results, signifying a balanced model performance regarding false positives and false negatives.

Several insights can be derived from the histograms:

- **Accuracy**: The distribution indicates how frequently certain accuracy values occur. A peak in the higher end of the range would signify that the model frequently achieves high accuracy.
- **ROC AUPR and ROC AUC**: These metrics provide insight into the model's performance over various threshold values. Peaks in their distributions can reveal the most common areas of performance for the model.

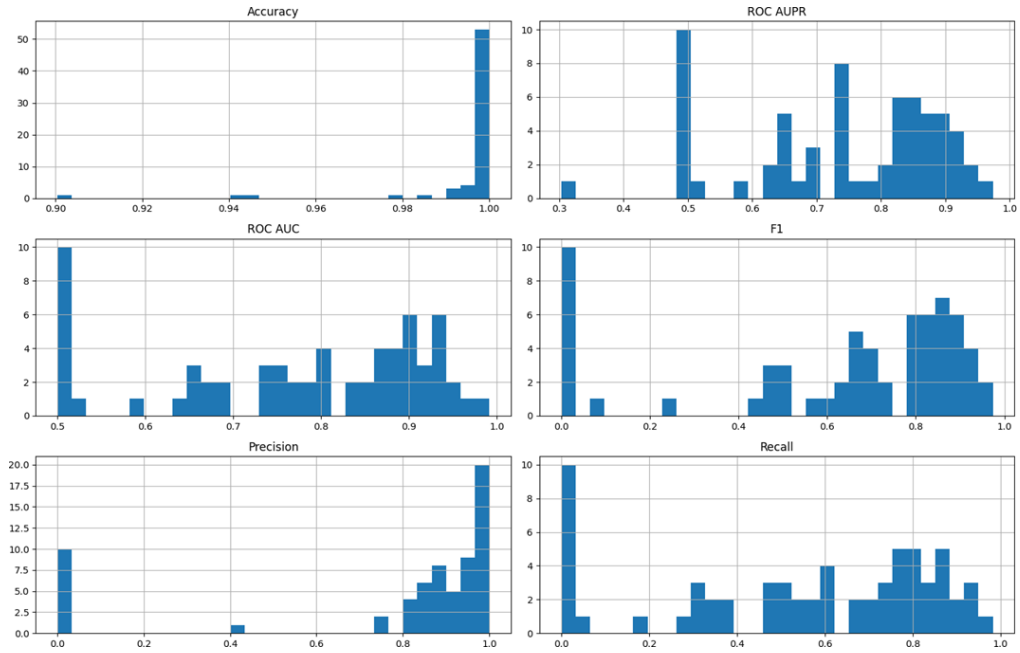

Fig. 3. Histograms showing distributions of model evaluation metrics

- **F1-score**: Observing the distribution of F1-score can provide a snapshot of the balance between precision and recall. A skew towards higher values suggests that the model maintains a consistent balance between false positives and false negatives.
- **Precision**: A histogram for precision can give insights into the model's capability to provide correct positive identifications. A higher frequency in the upper range indicates a model's strong capability to avoid false positives.
- **Recall**: The distribution of recall demonstrates the model's ability to correctly identify all actual positives. Peaks in the higher end would suggest that the model is adept at identifying most of the actual positive cases.

The histograms offer a visual representation of the model's performance across various evaluation metrics. They facilitate a quick comparative analysis between different metrics and allow for an understanding of the areas where the model excels or might need improvement.

C. Box Plots of Model Evaluation Metrics

The box plots provide a summarized visual representation of the distribution of the model evaluation metrics. In each box plot, the central rectangle spans the interquartile range, the segment inside the rectangle displays the median, and the whiskers extend to points that lie within 1.5 times the interquartile range of the lower and upper quartile

From the box plots as shown in Fig. 4, the following observations can be made:

- Accuracy: The box plot provides insights into the central tendency and variability of the accuracy values, as well as potential outliers.
- **ROC AUPR and ROC AUC**: The spread and central values of these metrics provide an overview of how the model performs across various thresholds.
- **F1-score**: The box plot for F1-score sheds light on its central value, spread, and any potential extremes in the distribution. This assists in understanding the model's balance between precision and recall.
- **Precision**: Through the box plot, one can discern the median value, variability, and potential outliers for precision, providing insights into the model's ability to make correct positive identifications.
- **Recall**: The box plot for recall aids in understanding the spread, central tendency, and potential extremes in the model's capability to identify all actual positives.

In conclusion, the box plots offer a compact and organized visualization of the model's performance metrics, highlighting their central tendencies, variability, and potential outliers. This aids in quickly gauging areas of strength and potential concern in the model's performance.

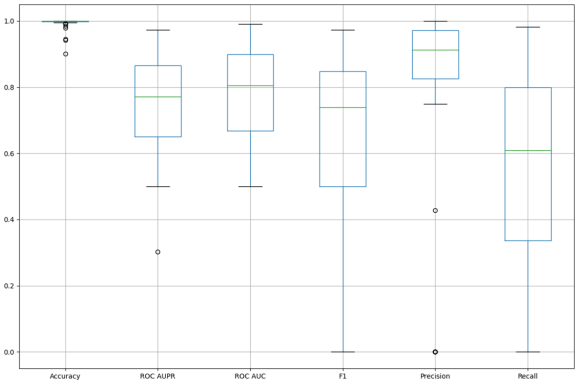

Fig. 4. Box plots showing distributions of model evaluation metrics

D. Top 10 Classes Analysis Based on Model Evaluation Metrics

The bar charts in Figs. 5-10 provide a visual representation of the top 10 classes across various model evaluation metrics, namely Accuracy, ROC AUPR, ROC AUC, F1-score, Precision, and Recall. The objective is to identify which classes perform exceptionally well for each metric.

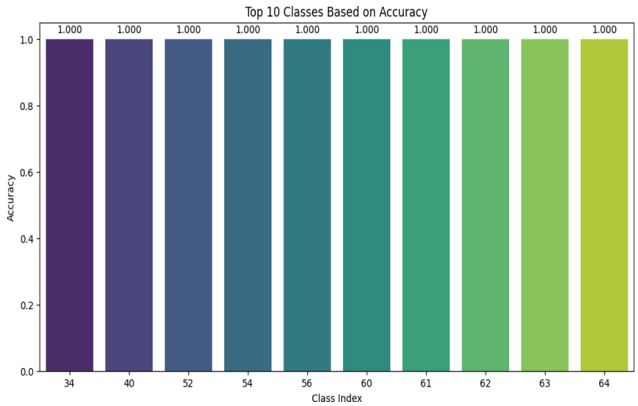

Fig. 5. Top 10 classes based on accuracy

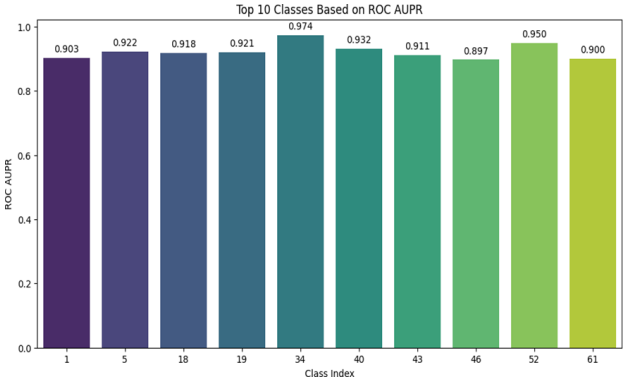

Fig. 6. Top 10 classes based on ROC AUPR

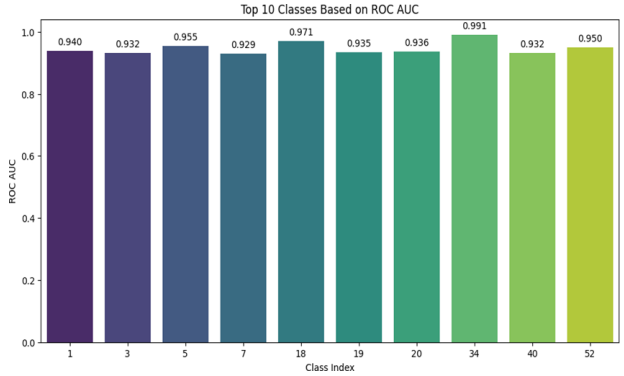

Fig. 7. Top 10 classes based on ROC AUC

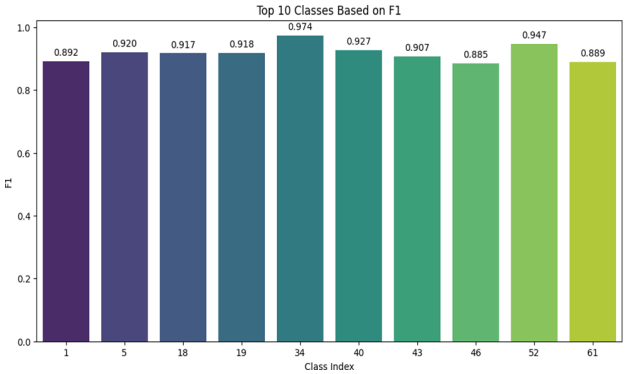

Fig. 8. Top 10 classes based on F1-score

From these bar charts, the following insights can be derived:

- **Accuracy**: The chart shows which classes achieve the highest accuracy scores, with the exact values annotated on each bar. This provides insights into which classes the model can predict with high confidence.
- **ROC AUPR and ROC AUC**: The top classes for these metrics indicate which categories the model performs optimally across various thresholds.
- **F1-score**: High F1-scores in the top classes denote a balanced harmonic mean between precision and recall for those classes.
- **Precision**: Top classes in this metric indicate the categories where the model can make correct positive identifications with higher confidence.
- **Recall**: The leading classes for recall indicate the categories in which the model is adept at identifying the majority of actual positives.

In conclusion, the bar charts of the top 10 classes for each evaluation metric offer valuable insights into the areas where the model excels. This information is pivotal when considering potential applications or further optimization of the model.

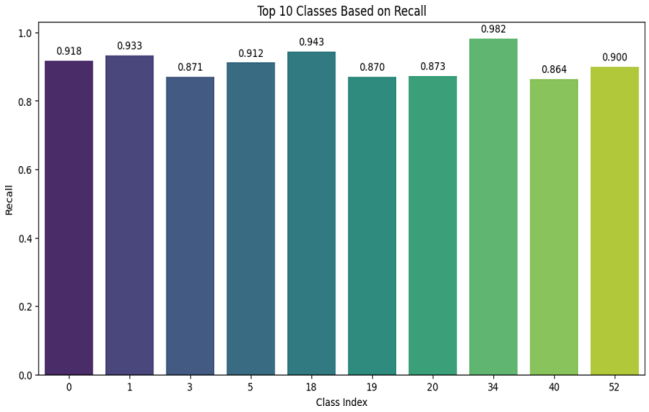

Fig. 9. Top 10 classes based on Recall

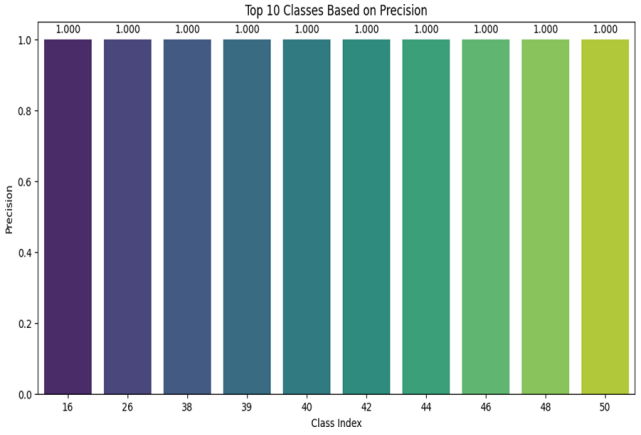

Fig. 10. Top 10 classes based on Precision

E. Overall Evaluation of Model Performance

The Table IV summarizes the performance metrics of the model on the dataset, as shown in Fig. 11. Both micro and macro averages are considered for several metrics, providing a comprehensive evaluation of the model's performance.

From the table, several important insights can be observed:

- The Accuracy of the model is 0.836518, which means the model correctly predicts 83.65% of the instances in the dataset.
- The Micro ROC AUPR and Macro ROC AUPR are 0.904623 and 0.829933, respectively. The Micro AUPR being higher suggests that the model performs better in terms of precision and recall on a global scale across all classes.
- The Micro and Macro ROC AUC values are close to 1, indicating excellent performance in distinguishing between the positive and negative classes across all thresholds.
- The Micro F1, Precision, and Recall scores are all the same, which means there's a balanced harmonic mean between precision and recall on a micro level.
- The Macro F1, Precision, and Recall scores are 0.621660, 0.775576, and 0.549186, respectively. These figures indicate that while the model's precision is relatively high, its recall on a macro level (averaging across classes) is somewhat lower.

  In summary, the model demonstrates strong predictive capabilities, especially in terms of the area under the ROC curve. However, the macro metrics indicate potential areas of improvement, especially in terms of recall across different classes.

TABLE IV.
SUMMARY OF THE DATASET

Metric	Value
Accuracy	0.836518
Micro ROC AUPR	0.904623
Macro ROC AUPR	0.829933
Micro ROC AUC	0.997406
Macro ROC AUC	0.987539
Micro F1	0.836518
Macro F1	0.621660
Micro Precision	0.836518
Macro Precision	0.775576
Micro Recall	0.836518
Macro Recall	0.549186

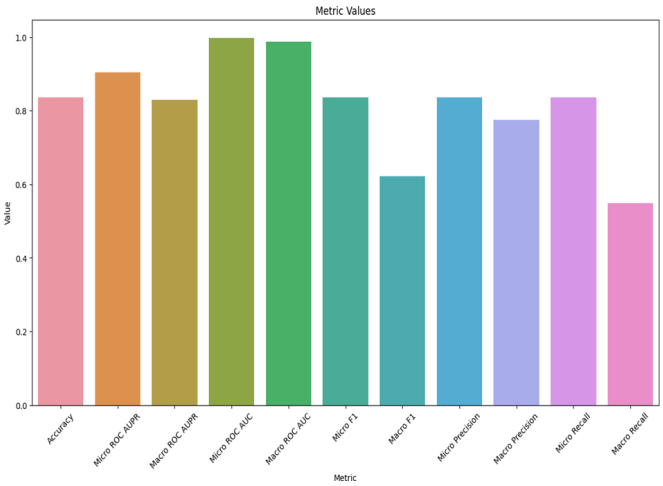

Fig. 11. Metric values

V. CONCLUSION

In this research, we presented a novel methodology for DDIs by synergizing attributed heterogeneous graph embedding with advanced deep learning mechanisms. Our approach seamlessly integrates a diverse range of drug attributes, capturing the intricacies of their interactions. The two-phase process not only aggregates data from various sources, including Chemical composition, Interaction Targets, Enzymatic reactions, and Biological Pathways, but also adeptly represents each drug within a multifaceted network, with the attributes serving as node descriptors. The inclusion of an attention-centric embedding mechanism accentuates the model's capability, ensuring the emphasis on vital drug interactions during predictions. By culminating this intricate process with a dense neural network, we achieved a more granular and insightful categorization of DDIs, setting a precedent for future works in drug interaction analysis.

FUTURE WORKS

As we move forward, there are several avenues to explore to further enhance our understanding and prediction of drug-drug interactions (DDIs). Firstly, the integration of additional datasets could supplement the existing attributes, thereby refining the drug representation within the heterogeneous network. This could potentially uncover less understood or latent drug interactions. Secondly, advancements in graph neural network architectures offer promising potential for better feature extraction and representation. We are keen on experimenting with more recent and innovative neural network architectures to see how they influence the accuracy and robustness of the model. Another intriguing prospect is the introduction of temporal dynamics, considering that drug properties and

interactions could evolve over time. Finally, a broader collaborative approach involving pharmacologists, clinicians, and data scientists could usher in a multidisciplinary perspective, which may lead to breakthrough findings in DDI predictions and patient-specific recommendations. The insights and feedback from this study will undeniably shape the trajectory of these future endeavors.

Conflict of Interest

The authors have no conflict of relevant interest to this article.

References

[1] K. Han, P. Cao, Y. Wang, F. Xie, J. Ma, M. Yu, J. Wang, Y. Xu, Y. Zhang, and J. Wan, "A review of approaches for predicting drug–drug interactions based on machine learning," *Frontiers in pharmacology*, vol. 12, p. 814858, 2022.

[2] Y. Tan, C. Kong, L. Yu, P. Li, C. Chen, X. Zheng, V. S. Hertzberg, and C. Yang, "4sdrug: Symptom-based set-to-set small and safe drug recommendation," in *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 3970–3980, 2022.

[3] Y. Qiu, Y. Zhang, Y. Deng, S. Liu, and W. Zhang, "A comprehensive review of computational methods for drug-drug interaction detection," *IEEE/ACM transactions on computational biology and bioinformatics*, vol. 19, no. 4, pp. 1968–1985, 2021.

[4] H.-C. Yi, Z.-H. You, D.-S. Huang, and C. K. Kwoh, "Graph representation learning in bioinformatics: trends, methods and applications," *Briefings in Bioinformatics*, vol. 23, no. 1, p. bbab340, 2022.

[5] M. Zitnik, M. Agrawal, and J. Leskovec, "Modeling polypharmacy side effects with graph convolutional networks," *Bioinformatics*, vol. 34, no. 13, pp. i457–i466, 2018.

[6] T. Takeda, M. Hao, T. Cheng, S. H. Bryant, and Y. Wang, "Predicting drug–drug interactions through drug structural similarities and interaction networks incorporating pharmacokinetics and pharmacodynamics knowledge," *Journal of cheminformatics*, vol. 9, no. 1, p. 16, 2017.

[7] K. Huang, C. Xiao, T. Hoang, L. Glass, and J. Sun, "Caster: Predicting drug interactions with chemical substructure representation," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 34, pp. 702–709, 2020.

[8] Y. Zheng, H. Peng, X. Zhang, Z. Zhao, X. Gao, and J. Li, "Ddi-pulearn: a positive-unlabeled learning method for large-scale prediction of drug-drug interactions," *BMC bioinformatics*, vol. 20, no. Suppl 19, p. 661, 2019.

[9] F. Tanvir, M. I. K. Islam, and E. Akbas, "Predicting drug-drug interactions using meta-path based similarities," in *2021 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, pp. 1–8, IEEE, 2021.

[10] Y.-H. Feng, S.-W. Zhang, and J.-Y. Shi, "Dpddi: a deep predictor for drug-drug interactions," *BMC bioinformatics*, vol. 21, no. 1, p. 419, 2020.

[11] Z.-H. Ren, C.-Q. Yu, L.-P. Li, Z.-H. You, Y.-J. Guan, X.-F. Wang, and J. Pan, "Biodkg–ddi: predicting drug–drug interactions based on drug knowledge graph fusing biochemical information," *Briefings in Functional Genomics*, vol. 21, no. 3, pp. 216–229, 2022.

[12] Y. Yu, K. Huang, C. Zhang, L. M. Glass, J. Sun, and C. Xiao, "Sumgnn: multi-typed drug interaction prediction via efficient knowledge graph summarization," *Bioinformatics*, vol. 37, no. 18, pp. 2988–2995, 2021.

[13] S. Mei and K. Zhang, "A machine learning framework for predicting drug–drug interactions," *Scientific Reports*, vol. 11, no. 1, p. 17619, 2021.

[14] S. Vilar, R. Harpaz, E. Uriarte, L. Santana, R. Rabadan, and C. Friedman, "Drug—drug interaction through molecular structure similarity analysis," *Journal of the American Medical Informatics Association*, vol. 19, no. 6, pp. 1066–1074, 2012.

[15] A. Gottlieb, G. Y. Stein, Y. Oron, E. Ruppin, and R. Sharan, "Indi: a computational framework for inferring drug interactions and their associated recommendations," *Molecular systems biology*, vol. 8, no. 1, p. 592, 2012.

[16] R. Ferdousi, R. Safdari, and Y. Omidi, "Computational prediction of drug-drug interactions based on drugs functional similarities," *Journal of biomedical informatics*, vol. 70, pp. 54–64, 2017.

[17] B. Davazdahemami and D. Delen, "A chronological pharmacovigilance network analytics approach for predicting adverse drug events," *Journal of the American Medical Informatics Association*, vol. 25, no. 10, pp. 1311–1321, 2018.

[18] H. Luo, P. Zhang, H. Huang, J. Huang, E. Kao, L. Shi, L. He, and L. Yang, "Ddi-cpi, a server that predicts drug–drug interactions through implementing the chemical–

protein interactome," *Nucleic acids research*, vol. 42, no. W1, pp. W46–W52, 2014.

[19] H. Ibrahim, A. M. El Kerdawy, A. Abdo, and A. S. Eldin, "Similarity-based machine learning framework for predicting safety signals of adverse drug–drug interactions," *Informatics in Medicine Unlocked*, vol. 26, p. 100699, 2021.

[20] A. Kastrin, P. Ferk, and B. Leskošek, "Predicting potential drug-drug interactions on topological and semantic similarity features using statistical learning," *PloS one*, vol. 13, no. 5, p. e0196865, 2018.

[21] X. Chen, X. Liu, and J. Wu, "Drug-drug interaction prediction with graph representation learning," in *2019 IEEE International conference on bioinformatics and biomedicine (BIBM)*, pp. 354–361, IEEE, 2019.

[22] T. Gaudelet, B. Day, A. R. Jamasb, J. Soman, C. Regep, G. Liu, J. B. Hayter, R. Vickers, C. Roberts, J. Tang, *et al.*, "Utilizing graph machine learning within drug discovery and development," *Briefings in bioinformatics*, vol. 22, no. 6, p. bbab159, 2021.

[23] K. M. Saifuddin, M. I. K. Islam, and E. Akbas, "Drug abuse detection in twitter-sphere: Graph-based approach," in *2021 IEEE International Conference on Big Data (Big Data)*, pp. 4136–4145, IEEE, 2021.

[24] Y. Bai, K. Gu, Y. Sun, and W. Wang, "Bi-level graph neural networks for drug-drug interaction prediction," *arXiv preprint arXiv:2006.14002*, 2020.

[25] B. Bumgardner, F. Tanvir, K. M. Saifuddin, and E. Akbas, "Drug-drug interaction prediction: a purely smiles based approach," in *2021 IEEE International Conference on Big Data (Big Data)*, pp. 5571–5579, IEEE, 2021.

[26] D. Weininger, "Smiles, a chemical language and information system. 1. introduction to methodology and encoding rules," *Journal of chemical information and computer sciences*, vol. 28, no. 1, pp. 31–36, 1988.

[27] J. Y. Ryu, H. U. Kim, and S. Y. Lee, "Deep learning improves prediction of drug–drug and drug–food interactions," *Proceedings of the national academy of sciences*, vol. 115, no. 18, pp. E4304–E4311, 2018.

[28] R. Gómez-Bombarelli, J. N. Wei, D. Duvenaud, J. M. Hernández-Lobato, B. Sánchez-Lengeling, D. Sheberla, J. Aguilera-Iparraguirre, T. D. Hirzel, R. P. Adams, and A. Aspuru-Guzik, "Automatic chemical design using a data-driven continuous representation of molecules," *ACS central science*, vol. 4, no. 2, pp. 268–276, 2018.