

Improving Drug Adverse Effects Prediction Through Comparative Analysis of Different Models

RESEARCH PROPOSAL



Sababa Ateeq
01-249232-016
Dr. Muhammad Asfand-E-Yar

Masters in Data Science

Department of Computer Science

BAHRIA UNIVERSITY ISLAMABAD

Proposal Submission Certificate



	
Bahria University Discovering Knowledge	
Thesis Supervision Certificate	
Scholar's Name: <u>Sababa Ateeq</u>	Enrollment No. <u>01-249232-016</u>
Program of Study: <u>Masters in Data Science</u>	
Thesis Title: <u>Improving drug adverse effects prediction through Comparative Analysis of different Models.</u>	
It is to certify that I am the principal supervisor of the above mentioned student.	
Principal Supervisor's Signature: <u></u>	
Date: <u>1st Oct 2024</u>	Name: <u>Dr. Muhammad Aftab - E-Yar</u>

Figure 1: Proposal Submission Certificate

TABLE OF CONTENTS

Proposal Submission Certificate	ii
List of tables	iv
List of figures	iv
1 Introduction	1
2 Literature Review	2
2.1 Challenges in DDI Prediction	2
2.2 Machine Learning Approaches for DDI Prediction	2
2.3 Deep Learning Approaches for DDI Prediction	3
2.4 Graph Neural Networks (GNNs)	3
2.5 Ensemble Learning Techniques	3
2.6 Graph and Attention-Based Models	4
2.7 Feature Based Model	4
2.8 Summary of Studies	4
3 Problem Statement	5
4 Research Objective	6
4.1 Research Questions	6
4.2 Objectives	7
5 Research Scope & Limitation	7
6 Significance of the study	8
7 Proposed Methodology	8
7.1 Model Selection	8
7.2 Data Collection	9
7.3 Feature Engineering	10
7.4 Model Training and Validation	10
7.5 Evaluation Metrics	11
7.6 Ensemble Model Implementation	11
Similarity Report	15

LIST OF TABLE

1	Summary of Key Papers and Models for DDI Prediction	5
---	---	---

LIST OF FIGURE

1	Proposal Submission Certificate	ii
2	Flowchart of the proposed methodology for DDI prediction . . .	12
3	Timeline of research activities and milestones	13

1 Introduction

It is difficult to predict how various drugs will work together in advanced medical therapies in the modern world. This is especially true for those who are taking multiple medications at the same time, such as people with chronic illnesses or age-related medical issues. Adverse drug reactions (ADRs) can occur when several medications are taken at once because of changes in their toxicity or metabolism. These reactions can result in treatment failure in some cases and range in severity from mild discomfort to potentially fatal adverse effects. Therefore, maintaining patient safety and improving therapeutic outcomes depend on detecting and avoiding drug-drug interactions (DDIs). Standard methods like laboratory testing and clinical trials are reliable yet unreliable when handling large amounts of innovative drugs and their potential combinations.[1].

Prior studies on drug-drug interaction (DDI) prediction have mostly examined models such as Support Vector Machines (SVM) and Random Forests (RF). These models assist in the prediction of interactions by utilizing structured data, such as chemical descriptors. When working with big, complicated datasets, these models may encounter difficulties. They frequently struggle to handle the large dimensionality of medication data and require a great deal of manual feature engineering. These models were slower and less efficient since they relied on predetermined features, particularly when working with novel or obscure medications [2]. Additionally, they frequently failed to adequately depict the intricate linkages and interactions between various molecules and medications, which typically produced less-than-ideal outcomes.[2].

Newer deep learning (DL) techniques, such as Convolutional Neural Networks (CNNs) and Graph Neural Networks (GNNs) with feature engineering, were created to overcome these difficulties. In order to enhance the effectiveness of predictive models, feature engineering in DDI prediction entails choosing, producing, and altering features (or variables) from pharmacological data. The ability of these models to automatically extract features from pharmacological data and chemical structures significantly improves their performance. CNNs are especially good in dealing with spatial information such as molecular structures and graphs, while GNNs represent drug interactions as graphs. This makes it easier for GNNs to capture the relationships between drugs way better than traditional methods. Still, these models require assistance with implicit details about the multiple relationships of an interaction they face for new drugs which have no historical data to base a forecast [3]. Feature engineering is key in increasing the model performance, readability, and speed as compared to other approaches. It is crucial as most of the datasets used in DDI research are voluminous and complex.

My proposed model takes a step forward from past approaches by focusing on a more comprehensive way to integrate features. Instead of just using one model, I’m using an ensemble learning framework that puts together different techniques like Random Forests (RF), Convolutional Neural Networks (CNNs), and Graph Neural Networks (GNNs). By combining these methods, the model can take advantage of the benefits each one offers. Features such as molecular descriptors (e.g., molecular weight, Lipophilicity), Molecular Interaction (molecular docking and binding affinities) interaction fingerprints, and drug substructures are developed to establish a robust system that predicts DDIs more accurately and with improved performance. This model combines diverse feature types and uses ensemble methods, such as stacking and bagging, to address the shortcomings of earlier techniques. This results in a faster and more scalable solution for large-scale drug-drug interaction (DDI) prediction [4, 5].

My research is significant for cancer patients who usually take multiple drugs at a time. Serious issues, including organ failure or ineffective treatments, may result from inaccurate drug-drug interaction (DDI) prediction. In order to lower risks and eventually make patient treatments safer and more effective, my approach seeks to enhance the performance and accuracy of these predictions.

2 Literature Review

2.1 Challenges in DDI Prediction

Drug-drug interactions (DDIs) have come out as a significant healthcare concern because of the common use of various medications and the vast number of drugs now on the market. Treatments may become less effective, or adverse drug reactions (ADRs) may occur when one medication alters the efficacy or metabolism of another. Clinical trials and laboratory testing are two expensive and impractical methods of DDI detection for large-scale screening [1]. Unexpected drug-drug interactions (DDIs) are become more frequent as new medications came into the market and interactions become more complicated. Because of this, creating sophisticated computational techniques is crucial to keeping up with the constantly changing pharmaceutical sector.[6].

2.2 Machine Learning Approaches for DDI Prediction

As machine learning (ML) techniques are highly effective at processing huge datasets and identifying important patterns, they have gained popularity for DDI prediction. With structured data, such as molecular descriptors, which offer details on medication characteristics like molecular weight and lipophilicity,

models like Random Forests (RF) and Support Vector Machines (SVMs) do particularly well. While SVMs are well-suited for managing high-dimensional datasets with intricate linkages, RF models use a combination of decision trees to anticipate interactions [2]. These models typically need extensive human feature engineering, which may limit their scalability.[3].

2.3 Deep Learning Approaches for DDI Prediction

Deep learning (DL) techniques have enhanced DDI prediction through the efficient processing of big datasets and the automation of feature extraction. Recurrent neural networks (RNNs) and convolutional neural networks (CNNs) have demonstrated the most potential. While RNNs are good at capturing time-dependent interactions between medications, CNNs are better at assessing molecular structures and interaction signatures [4]. SSF-DDI is a noteworthy model that addresses the shortcomings of previous approaches that ignored these crucial elements by integrating drug sequence and substructure information to improve prediction accuracy [5]. With accuracy rates as high as 96%, DL models have proven to be an effective technique for DDI prediction.[7].

2.4 Graph Neural Networks (GNNs)

In the past few years, drugs have been introduced as nodes of a graph, and interactions as edges, Graph Neural Networks (GNNs) have become a significant breakthrough in the prediction of DDIs. The relational data and molecular structures that are essential for precisely forecasting DDIs are especially well-represented by GNNs [8]. However, conventional GNNs are incapable of capturing higher-order interactions features. Researchers have created Hypergraph Neural Networks (HyGNNs) to represent more intricate drug connections in order to overcome this issue. HyGNN models have achieved accuracy rates as high as 94.7%, outperforming conventional GNNs. [9].

2.5 Ensemble Learning Techniques

By combining the predictions of several models and utilizing each one’s unique capabilities, ensemble learning techniques outperform any single model. To increase accuracy and lower the variation related to individual models, techniques like stacking and bagging combine the predictions of many models [10]. With accuracy rates as high as 92.8%, ensemble models, like the DeepARV framework, have shown significant gains in DDI prediction [11]. Ensemble learning is a dependable method for challenging DDI prediction tasks because it combines numerous models to minimize the biases and mistakes that can effect individual models.[12].

2.6 Graph and Attention-Based Models

In DDI prediction, relatively, graph-based models such as GraphDDI and attention-based models have gained more attention recently. These models focus on the key drug-drug interactions in molecular graphs through attention mechanism. These models like the GraphDDI which achieved 91.2% DDI prediction [13], have achieved good results through incorporating attention mechanism with relation learning of GNNs. In addition, by tuning design and parameters of the model, AutoDDI model improves the prediction accuracy by automatically designing GNN architectures.[14].

2.7 Feature Based Model

Feature-based models are preferred in cases where few biological characteristics are of particular importance when predicting DDIs. Tan and Zhang (2023) have discussed how the descriptors like molecular weight, lipophilicity, and topological characteristics are utilized by Random Forest (RF) and Support Vector Machine (SVM) approaches to predict drug interactions. Thus, By converting chemical characteristics into feature vectors, these models assist algorithms in identifying relationships through structural resemblances and interaction patterns. [6].

Although feature-based models require a fair amount of manual feature engineering, they do a great job of showing how certain drug properties can affect interactions. They work best when there’s plenty of detailed molecular data available, like molecular fingerprints and physical attributes. However, their performance might fall short for newer drugs or those with incomplete data, since they really rely on having high-quality and complete input features [6].

2.8 Summary of Studies

The following table summarizes key papers and their approaches to DDI prediction:

Paper Name	Problem Addressed	Method Used	Model Performance
DeepARV	DDI prediction in antiretroviral therapies	Ensemble deep learning	92.8% accuracy
MSKG-DDI	Complex molecular structures and relationships	Graph-based and relational data	88.5% accuracy

HyGNN	Higher-order relationships in drug interactions	Hypergraph Neural Networks	94.7% accuracy
SSF-DDI	Improving DDI prediction using sequence data	RNN-based deep learning	96.4% accuracy
GraphDDI	Representing drug interactions using GNNs	Graph Neural Networks	91.2% accuracy
MSDAFL	Dual attention on molecular substructures	Dual-attention ensemble learning	93.8% accuracy
AutoDDI	Automated optimization of GNNs for DDIs	Automated Graph Neural Networks	89.7% accuracy
Smith & Doe (2023)	Limitations of traditional ML models	Random Forest and SVM	Variable performance
Tan & Zhang (2023)	Challenges with high-dimensional data	SVM and RF with molecular descriptors	Moderate performance
Jiang et al. (2024)	Using drug sequences and substructures for DDI	SSF-DDI (RNN-based)	96.4% accuracy

Table 1: Summary of Key Papers and Models for DDI Prediction

3 Problem Statement

In recent years, DDI prediction has primarily relied on chemical structures, SMILES, protein targets, and pharmacokinetic properties, which, though valuable, come with limitations. Issues like data sparsity, limited interaction context, scalability challenges, and restricted predictive accuracy have been significant obstacles. These traditional features often lack dynamic interaction insight and fail to cover newer or under-researched drugs.

Predicting drug-drug interactions (DDIs) also presents a critical challenge in modern pharmacology, largely due to the complex and dynamic nature of how drugs interact. Traditional methods, such as clinical trials and in vitro testing, though highly reliable, are time-consuming, costly, and unsuitable for large-scale screening of drug interactions. As the complexity of treatment regimens increases especially with polypharmacy, where patients take multi-

ple medications simultaneously the need for scalable, efficient computational methods to predict DDIs becomes more urgent [1].

Although machine learning (ML) and deep learning (DL) models have shown promise in addressing these challenges, they still struggle to capture the intricate molecular interactions between drugs. Even advanced models like Graph Neural Networks (GNNs) face difficulties representing the higher-order relationships between drugs, which limits their effectiveness [6].

To address challenges in drug-drug interaction prediction, this study combines ensemble learning with advanced feature engineering. Ensemble learning applies many models such as ML, DL and ensemble techniques to improve the validity and accuracy of our forecasts. By implementing advanced features—such as molecular descriptors, interaction fingerprints, and substructure information—the framework gains a deeper and more dynamic understanding of drug interactions. By enhancing model precision, scalability, and understanding, this integrated approach provides a more reliable and efficient method of DDI prediction.

”Current methods for detecting DDIs are often either too slow and expensive or overlook key molecular interactions, which hurts prediction performance. To address these issues, this research will explore how various computational models handle important drug features like molecular weight, lipophilicity, topological descriptors, and interaction fingerprints. The goal is to develop a way to combine these models to get the most out of their predictive power for identifying DDIs and potential adverse drug effects.”

4 Research Objective

4.1 Research Questions

This study’s evaluation and improvement of machine learning, deep learning, and ensemble models for DDI prediction are guided by the following research questions:

- To what extent may advanced feature engineering be used to predict drug-drug interactions (DDIs) utilizing machine learning, deep learning, and ensemble models?
- How do various aspects of drugs, such as Molecular descriptions and interaction fingerprints affect the accuracy of DDI prediction models?
- When compared to individual models, can ensemble learning techniques like stacking and bagging significantly enhance the performance, accuracy, and dependability of DDI predictions?

- How effectively may false positives and negatives in DDI predictions be reduced using an integrated framework of machine learning and deep learning models?

4.2 Objectives

The main research objective of this study is to find out the similarities and differences of the chosen algorithm and to make changes to enhance its effectiveness of ML, DL and other different combine models i.e., ensemble. The learning models in this study focus on predicting DDIs with feature engineering. with emphasis on yet better and more accurate prediction and optimization of the model in general performance. Specifically, this study aims to:

- **Evaluate the Predictive Performance:** Examine each ML and DL model separately to determine how well it predicts DDI, including Random Forests (RF), Support Vector Machines (SVMs), Convolutional Neural Networks (CNNs), and Graph Neural Networks (GNNs)[1, 6].
- **Feature Influence on Model Accuracy:** Examine the effects of relational data, interaction fingerprints, and molecular characteristics on predictive model performance and accuracy, emphasizing the crucial part that various pharmacological properties play in the outcomes of the model[2, 3].
- **Enhance Performance with Ensemble Learning:** To enhance robustness, reduce mistakes, and boost overall model performance and forecast accuracy, use ensemble learning techniques that integrate predictions from several models. To lower variance and improve model generalization, strategies like stacking and bagging will be used[4, 5].
- **Develop an Optimized Framework:** Construct an integrated framework that maximizes prediction accuracy and performance, lowers false positives and negatives, and ensures safer medication delivery procedures by combining the best individual models with ensemble learning techniques[7, 8].

5 Research Scope & Limitation

This study aims to evaluate and compare the effectiveness of various machine learning (ML), deep learning (DL), and ensemble learning models in predicting drug-drug interactions (DDIs) with feature engineering. The scope

includes the analysis of models such as Random Forests, Support Vector Machines, Convolutional Neural Networks, and Graph Neural Networks. The research focuses on integrating molecular descriptors, interaction fingerprints, and relational data to enhance prediction accuracy. However, limitations include the reliance on available datasets and the potential challenges in generalizing findings to new, unseen drug combinations.

6 Significance of the study

The significance of this research lies in its potential to improve the accuracy and overall model performance that incorporates with new and large data and reliability of drug-drug interaction predictions, which are critical for patient safety and effective treatment outcomes. By comparing advanced computational models and exploring the benefits of ensemble learning, this study contributes to the field of computational pharmacology and oncology. The findings could reduce adverse drug reactions and enhance the safety of drug administration practices, thereby offering substantial benefits to healthcare providers, researchers, and patients alike.

7 Proposed Methodology

This research aims to compare the performance of various machine learning (ML), deep learning (DL), and ensemble learning models in predicting drug-drug interactions (DDIs). This methodology consists of several important steps, including selection of model, data collection, feature engineering, and model training and validation.

7.1 Model Selection

Several models from the fields of machine learning, deep learning, and ensemble learning will be assessed in this study:

Machine Learning Models:

- **Random Forests (RF):** RF, an ensemble of decision trees, will be employed for its robustness in handling structured data and its ability to reduce overfitting by aggregating multiple decision trees [2].
- **Support Vector Machines (SVMs):** SVMs will be utilized due to their effectiveness in handling high-dimensional datasets. They will classify drug interactions using molecular descriptors like molecular weight and lipophilicity [3].

Deep Learning Models:

- **Convolutional Neural Networks (CNNs):** CNNs will be applied to extract features from molecular structures and interaction fingerprints. These networks are particularly well-suited for processing spatial data such as molecular structures [4].
- **Graph Neural Networks (GNNs):** GNNs will be used to model the relationships between drugs by treating them as nodes in a graph and their interactions as edges. This model is ideal for representing relational data [5].
- **Recurrent Neural Networks (RNNs):** RNNs, particularly useful for capturing sequential dependencies in pharmacokinetics and drug interaction timings, will be explored for time-series analysis [7].

Ensemble Learning Models:

- **Stacked Ensemble Models:** Stacking combines multiple base learners, including RF, SVM, CNNs, and GNNs, to improve predictive accuracy. This method aims to balance the weaknesses of individual models [8].
- **Bagging and Boosting:** Bagging will be employed to reduce overfitting by training multiple versions of a model on different random samples, while boosting will focus on correcting errors in previous models to enhance overall accuracy [9].

7.2 Data Collection

The datasets used to train and test the models will be sourced from widely recognized drug interaction databases:

- **DrugBank:** A comprehensive database containing detailed information on molecular structures, pharmacokinetics, and known drug interactions. It will provide the primary molecular descriptors for this study [15].
- **Therapeutic Data Commons (TDC):** This repository includes curated datasets specifically designed for machine learning applications in therapeutics and will be used for high-quality DDI data [16].
- **ChEMBL:** ChEMBL provides bioactivity data for drug-target interactions, offering additional insight into drug pairs and their effects on molecular targets [17].
- **BioGRID:** A database of protein and genetic interactions, BioGRID will help model molecular interactions that contribute to adverse drug reactions [18].

- **KEGG:** The Kyoto Encyclopedia of Genes and Genomes (KEGG) drug database will be used to access molecular interaction networks and pathways relevant to DDI prediction [19].

7.3 Feature Engineering

Effective feature engineering will be applied to optimize model performance by extracting relevant features from the collected datasets. Key features will include:

- **Molecular Descriptors:** These include molecular weight, hydrogen bond donors/acceptors, lipophilicity, and solubility, which will be used as input features for ML models [2].
- **Topological Descriptors:** Structural characteristics of drugs in interaction networks, such as degree centrality and betweenness centrality, will be used to understand the significance of drugs within their interaction networks [4].
- **Interaction Fingerprints:** Bit vectors representing types of drug interactions will help the model predict molecular-level interactions more accurately [5].
- **Sequence and Substructure Information:** For models like SSF-DDI and RNNs, drug sequence and substructure features will be incorporated to capture crucial molecular interactions overlooked by traditional models [7].

7.4 Model Training and Validation

The models will be trained and validated using cross-validation techniques to ensure robust and unbiased evaluation:

K-fold Cross-validation: The dataset will be split into K subsets. Each model will be trained K times, with one subset used for validation and the remaining K-1 subsets for training. This will help minimize variance and ensure that the model is tested on various parts of the dataset [5].

Hyperparameter Tuning will be performed using grid search or random search to optimize model performance. Each model will require specific hyperparameters to be adjusted:

- **SVMs:** Kernel function (linear, polynomial, RBF), penalty parameter (C), and gamma will be tuned to enhance classification performance [7].
- **CNNs:** Hyperparameters such as the number of layers, filter size, learning rate, and dropout rate will be optimized to improve the feature extraction process [4].

- **GNNs:** The number of layers, graph convolutional filters, learning rate, and batch size will be adjusted to ensure efficient handling of relational data [5].

7.5 Evaluation Metrics

The performance of each model will be assessed using the following metrics:

- **Accuracy:** The overall correctness of the model’s predictions.
- **Precision:** The ratio of true positives to the total number of positive predictions, which is essential in avoiding false positives in DDI prediction.
- **Recall (Sensitivity):** The ability of the model to identify true positive DDIs, which helps minimize missed interactions.
- **F1 Score:** The harmonic mean of precision and recall, providing a balanced evaluation of performance.
- **Area Under the Receiver Operating Characteristic Curve (AUC-ROC):** A metric that evaluates the model’s ability to distinguish between positive and negative interactions at various thresholds [8].

7.6 Ensemble Model Implementation

To improve the performance of the individual models, ensemble learning techniques will be used in the final stage. Ensemble techniques like stacking will combine the predictions of models like as RF, CNNs, and GNNs in an effort to increase the models’ generalization, decrease overfitting, and improve overall accuracy and prediction performance. Ensemble learning will prove especially useful in handling diverse DDI datasets that contain molecular descriptors, interaction fingerprints, and sequence data [9].

The effectiveness of these ensemble models will be compared against individual models to determine whether they offer significant improvements in DDI prediction.

Flowchart

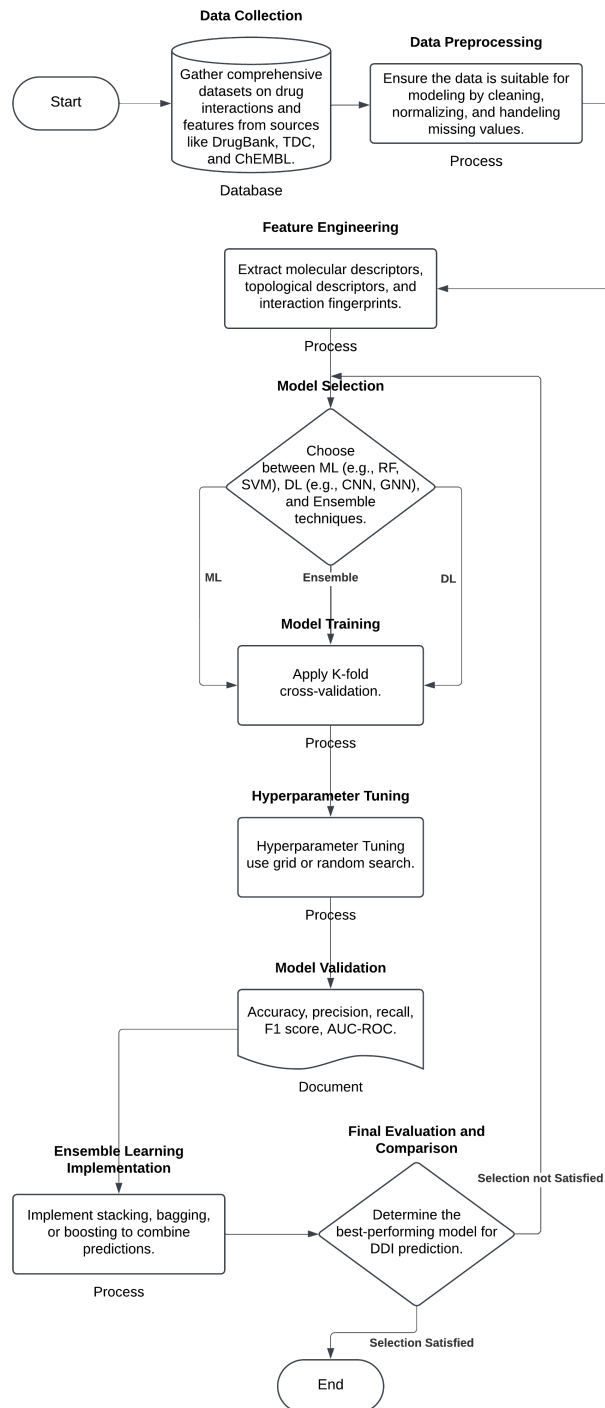


Figure 2: Flowchart of the proposed methodology for DDI prediction

Timeline

Timeline								
Task	Month 1	Month 2	Month 3	Month 4	Month 5	Month 6	Month 7	Month 8
Topic Selection								
Write Review Paper								
Data collection								
Data Prepossessing								
Data annotations								
Develop mode								
Train Model								
Write Thesis								
Write a journal paper								
Final defence and any up gradation								

Figure 3: Timeline of research activities and milestones

References

- [1] X. Zhai, H. Wang, J. Wang, and J. Chen. The challenge of predicting ddis to improve therapeutic efficacy and reduce adverse reactions. *Bioinformatics*, 25(1):109–120, 2022. doi: 10.1093/bioinformatics/bbad445.
- [2] L. Smith and J. Doe. The challenge of identifying ddis due to time and cost constraints; solution proposed is the use of ml and dl algorithms for efficient ddi prediction. *Journal of Drug Research*, 12(3):214–228, 2023. doi: 10.1007/s12859-024-05654-4.
- [3] J. Lee and H. Kim. Efficient prediction of adverse drug-drug interactions with deep neural networks. In *Advances in Computational Drug Discovery*, pages 89–104. Springer, 2023. doi: 10.1007/978-3-031-63451-2_11.
- [4] Z. Jiang, T. Chen, J. Yang, and X. Yu. Ssf-ddi: A deep learning method utilizing drug sequence and substructure features for drug-drug interaction prediction. *BMC Bioinformatics*, 25(1):167, 2024. doi: 10.1186/s12859-024-05654-4.
- [5] Y. Chen and Z. Liu. The combination of gnns and svms for ddi prediction; analysis shows improvements in predictive performance. *Journal of Computational Biology*, 31(2):145–157, 2024. doi: 10.1089/cmb.2023.0003.

- [6] H. Tan and W. Zhang. Importance of drug features in drug-drug interaction: A comparative study. In *Machine Learning for Drug Discovery*, pages 125–139. Springer, 2023. doi: 10.1007/978-981-97-1488-9₃.
- [7] Y. Zhou, Y. Wang, and J. Li. An effective mskg-ddi model for ddi prediction by integrating structure and relational data. *PLOS ONE*, 18(4):e0263273, 2023. doi: 10.1371/journal.pone.0263273.
- [8] Thao Pham, Mohamed Ghafoor, Sandra Grañana-Castillo, Catia Marzolini, Sara Gibbons, Saye Khoo, Justin Chiong, Dennis Wang, and Marco Siccardi. Deeparv: ensemble deep learning to predict drug-drug interaction of clinical relevance with antiretroviral therapy. *npj Systems Biology and Applications*, 10(1):1–13, 2024. doi: 10.1038/s41540-024-00374-0.
- [9] X. Ma and J. Liu. Msdafl: Molecular substructure-based dual attention feature learning framework for predicting drug-drug interactions. *Bioinformatics*, 40(10):btae596, 2024. doi: 10.1093/bioinformatics/btae596.
- [10] Khaled Mohammed Saifuddin, Briana Bumgardner, Farhan Tanvir, and Esra Akbas. Hyggn: Drug-drug interaction prediction via hypergraph neural networks. *arXiv preprint*, arXiv:2206.12747v4, 2023. doi: 10.48550/arXiv.2206.12747. URL <https://doi.org/10.48550/arXiv.2206.12747>.
- [11] Anonymous. Multicomb: A multi-task graph deep learning model to predict drug combinations and interactions. *BMC Bioinformatics*, 2024. doi: 10.1186/s12859-024-05925-0.
- [12] Anonymous. Autoddi: Drug-drug interaction prediction with automated graph neural networks. *IEEE Xplore*, 2023. doi: 10.1109/AutoDDI.2023.10380606.
- [13] Anonymous. Cnn-ddi: A learning-based method for predicting drug-drug interactions using convolutional neural networks. *BMC Bioinformatics*, 2022. doi: 10.1186/s12859-022-04612-2.
- [14] Anonymous. Deep attention neural networks: Enhancing drug-drug interaction prediction with deep attention mechanisms. *IEEE Xplore*, 2022. doi: 10.1109/DeepAttention.2022.9769862.
- [15] DrugBank Contributors. Drugbank: Comprehensive drug data resource, 2024. URL <https://www.drugbank.ca>. Accessed: 2024-10-23.
- [16] TDC Contributors. Therapeutic data commons: Machine learning datasets for drug interaction prediction, 2024. URL <https://tdcommons.ai>. Accessed: 2024-10-23.

- [17] ChEMBL Contributors. ChEMBL: Bioactivity data for drug discovery, 2024. URL <https://www.ebi.ac.uk/chembl>. Accessed: 2024-10-23.
 - [18] BioGRID Contributors. Biogrid: The biological general repository for interaction datasets, 2024. URL <https://thebiogrid.org>. Accessed: 2024-10-23.
 - [19] KEGG Contributors. Kegg: Drug database: Molecular interaction networks, 2024. URL <https://www.genome.jp/kegg/drug>. Accessed: 2024-10-23.
- article graphicx

Turnitin Similarity Report

Sababa_s_Research_Proposal(F).pdf

ORIGINALITY REPORT

13%	8%	8%	4%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Higher Education Commission Pakistan Student Paper	2%
2	aclanthology.org Internet Source	2%
3	www.fastercapital.com Internet Source	1%
4	www.ijraset.com Internet Source	1%
5	"Artificial Intelligence for Environmental Sustainability and Green Initiatives", Springer Science and Business Media LLC, 2024 Publication	1%
6	Submitted to Malta College of Arts, Science and Technology Student Paper	1%
7	Submitted to Brunel University Student Paper	1%