ACS OMEGA

Open Access

http://pubs.acs.org/journal/acsodf

Review

A Systematic Review of Drug-Related Interactions—Utilizing Deep Learning and LLMs for Prediction and Mitigation

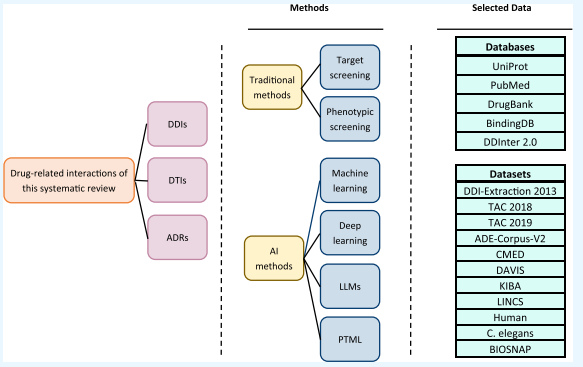
Soheila Behrooznia and Mohsen Hooshmand*

Cite This: *ACS Omega* 2025, 10, 61048−61075

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Computational drug discovery is essential for screening potential treatments and reducing the costs and time associated with proposing or combining drugs for disease management. Despite the extensive research conducted in this field, it remains an emerging area, particularly with the advent of machine learning, deep learning, and large language models (LLMs). This systematic review examines the integration of machine learning and deep learning techniques in drug discovery, concentrating on three critical areas: drug−drug interactions (DDIs), drug-target interactions (DTIs), and adverse drug reactions (ADRs). The review analyzes over 100 papers published between 2020 and 2025, categorizing the methods into deep learning, machine learning, graph learning, and hybrid models. It highlights the transformative impact of natural language processing (NLP) and LLMs in extracting meaningful insights from biomedical literature and chemical data. Furthermore, this work introduces key databases and data sets widely utilized in drug discovery. Additionally, this review identifies gaps in the existing research, such as the lack of comprehensive studies that simultaneously address DDI, DTI, and ADR extraction, and it proposes a more holistic approach to fill these gaps. The paper concludes by thoroughly evaluating various models, underscoring their performance metrics.



1. INTRODUCTION

Drug discovery (DD), as a systematic scientific process, aims to identify, design, and develop new therapeutic agents to cure or prevent diseases and improve medical conditions. The term "pipeline" is frequently used to describe the unidirectional progression from the hit or the lead to the candidate and the marketed drug, which is supported by both scientific and clinical research. In fact, the process is iterative in nature, diverse, and complicated. It has several steps, including target identification, lead optimization, clinical trials, regulatory assessment, and postmarketing surveillance. Several techniques are used to support the process, i.e., computational modeling, high-throughput screening, and biological testing.[1] Traditional methods require a large amount of clinical tests, whereas novel methods rely on modeling and bioinformatic techniques that accelerate the discovery of new drugs.[2,3] In addition, drug discovery requires a significant investment in finance and expert human resources.[4]

Traditional drug discovery involves screening methods, such as target-based or phenotypic screening. In target-based screening, compounds are selected for their ability to bind to a biological target, while phenotypic screening tests compounds against an organism without prior knowledge of the target.[5−8] These methods are complex, time-consuming, and costly.[9]

Drug research requires understanding disease biology at the cellular and molecular level, identifying drug targets, and selecting chemical molecules for testing. These molecules undergo efficacy and safety assessments in preclinical phases before clinical trials.[10] Bringing a new drug to market takes 10−15 years and costs 800M−1.5B, with recent estimates reaching 2.5B dollars.[11] However, the drug development process faces high failure rates due to toxicity and side effects.[4]

For over three decades, computer-aided DD/design (CADD) has been essential in developing small therapeutically relevant compounds.[12,13] CADD is divided into structure-based, ligand-based, and systematic-based approaches.[14,15]

Structure-based approaches, similar to high-throughput screening, require structural data of both the target and the ligand. Researchers study proteins or biomarkers related to pathophysiology and use 3D target protein structures to calculate interaction energies.[15] Key methods include DMAP, molecular docking, and molecular dynamics simulations.[16−18] Ligand-based approaches estimate activity based on ligand similarities when 3D protein structures are unavailable.[14] They focus on identifying features in chemical compounds that influence biological activity.[15]

Received: May 27, 2025
Revised: November 8, 2025
Accepted: December 1, 2025
Published: December 11, 2025

Systematic-based approaches, such as Proteochemometric Modeling, view biological data holistically to assess how chemical compounds affect cellular interactions, aiding in drug repurposing.[14,19]

The integration of data science, informatics, and AI has revolutionized drug discovery, improving efficiency and reducing costs.[20,21] AI, particularly machine learning and deep learning, is increasingly used to handle large, complex data sets, making drug research more efficient.[20,22−24] These technologies automate the drug discovery pipeline, especially in preclinical stages, enhancing productivity and lowering costs.[1,4,25]

AI accelerates drug discovery by improving target identification, drug screening, and lead optimization.[26] Traditional methods are labor-intensive and time-consuming, but AI efficiently analyzes genomic and protein data to identify drug targets.[20] Machine learning models process complex biological data and predict DTIs, aiding in target discovery. AI also predicts bioactivity and molecular properties, such as toxicity and ADME profiles, reducing the need for expensive screening methods.[25] Deep learning offers superior accuracy in predicting molecular properties, while AI tools like reinforcement learning and generative models optimize lead compounds, enhancing pharmacological properties and trial success rates.[25]

Traditional approaches that focus on only one target often struggle to treat complex conditions like cancer and neurodegenerative diseases. With more biological data and better algorithms, machine learning has become a useful tool to help develop drugs that work on multiple targets. Recently, the method of discovering drugs is moving away from the old idea of "one drug, one target" to a more complete system approach called multitarget DD. This change is because many diseases involve multiple genes, proteins, and pathways, so targeting just one is often insufficient. Designing drugs that hit several targets at once can improve treatment, decrease resistance, and be safer. This idea is part of systems pharmacology, which uses network biology, pharmacokinetics, and computer models to understand how drugs work in the body.[27] PTML, or Perturbation-Theory Machine Learning models, can predict multiple biological effects (such as activity, toxicity, and pharmacokinetics) simultaneously against different targets, including proteins, microbial strains, cell lines, laboratory animals, and humans, among others. It also considers information from various test conditions. Therefore, PTML modeling has been used in the discovery and design of molecules to fight microbial infections, neurological disorders, nanosystems for drug release or treatment, and various areas at the border of immunology and toxicology.[28] Therefore, PTML is a sophisticated computational modeling approach that combines perturbation theory principles with machine learning algorithms to analyze complex biological networks and molecular data.[29]

PTML models utilize perturbation operators derived from Markov linear indices to investigate local and global variations in Metabolic Reaction Networks (MRNs) across different organisms, which enables accurate prediction of network connectivity patterns and validation of metabolic models with high precision.[29] PTML integrates chemical descriptors, such as fragment-based topological indices and physicochemical properties, with experimental conditions to develop interpretable models capable of predicting biological activities and designing compounds with specific multitarget effects, including dual inhibitors for therapeutic applications.[30,31] Furthermore, PTML leverages graph-based descriptors and machine learning techniques like multilayer perceptrons (MLPs) to understand how molecular structures and experimental variations influence biological responses. This framework facilitates virtual screening and the in silico development of bioactive molecules, exemplified by its application in the design of versatile anticancer agents.[31] Overall, PTML provides a versatile platform that combines perturbation theory, structural descriptors, and machine learning to predict and optimize biological and chemical properties, advancing DD and systems biology research.[29−31]

Some work has been done on this topic: The model proposed by Kleandrova and Speck-Planche[32] integrates PTML via MLPs to predict multiprotein and multicell inhibitors targeting pancreatic cancer (PANC), using molecular descriptors from topological indices, physicochemical properties, and atom-based quadratic indices, achieving over 78% accuracy in classifying active compounds and enabling rational therapeutic design. Another work developed a PTML-MLP model predicting dual NET and SERT inhibitors from ChEMBL data, demonstrating high accuracy, robustness, and key structural features, such as reduced hydrophobicity and specific aromatic rings that guide DD for mood disorders.[31] Also, Diéguez-Santana et al.[29] introduced the CPTML model, which generalizes atom-based linear indices to Markov linear indices for analyzing complex metabolic networks across different organisms, including random and real networks. In addition, PTML was used to identify chemicals, like Linagliptin, Icariin, and Rolipram, as potential dual inhibitors of key inflammatory proteins, making it useful for developing broad anti-inflammatory treatments, including for COVID-19.[33] Another work included a multitask PTML model for predicting antibacterial activity against Gram-negative pathogens and safety profiles, which employed linear discriminant analysis and molecular fragments; an MLP-based PTML model for designing inhibitors against pathogenic parasites; a PTML model for virtual screening of antituberculosis agents targeting multiple strains; and a PTML model for designing drugs with pan-antiviral and anticytokine storm properties utilizing artificial neural networks, all integrating molecular topology with machine learning for multiend point and multitarget predictions in anti-infective DD.[34]

In summary, AI has transformed drug discovery by improving various stages through computational models such as RNNs,[35−37] VAEs,[38] GNNs,[39] transformers,[40,41] and MPNNs.[42,43] These models enable the generation of novel molecules by exploring chemical spaces based on properties like toxicity and binding affinity.[35−38,44] They also facilitate accurate predictions of molecular attributes and chemical reactions, improving drug screening and lead optimization via data analysis and pattern recognition.[40−43] AI's impact extends to drug target identification, clinical trial optimization, and personalized medicine, cementing its role in modern therapeutic development.[25]

NLP plays a key role in computational drug discovery, transforming both human language modeling and chemistry, which is central to drug development.[45,46] Just as human language uses tokens (words and subwords), chemistry is represented by sequences of amino acids, atoms, or chemical reactions,[46] enabling the application of NLP techniques like transformer models in chemical research.[47−50]

Transformers excel at capturing contextual relationships, making them effective for tasks like translation and summarization.[51] In chemistry, they are critical for protein structure prediction (e.g., AlphaFold2[52,53]) and reaction outcome prediction.[46] The integration of multimodal data, including

https://doi.org/10.1021/acsomega.5c04997
ACS Omega 2025, 10, 61048−61075

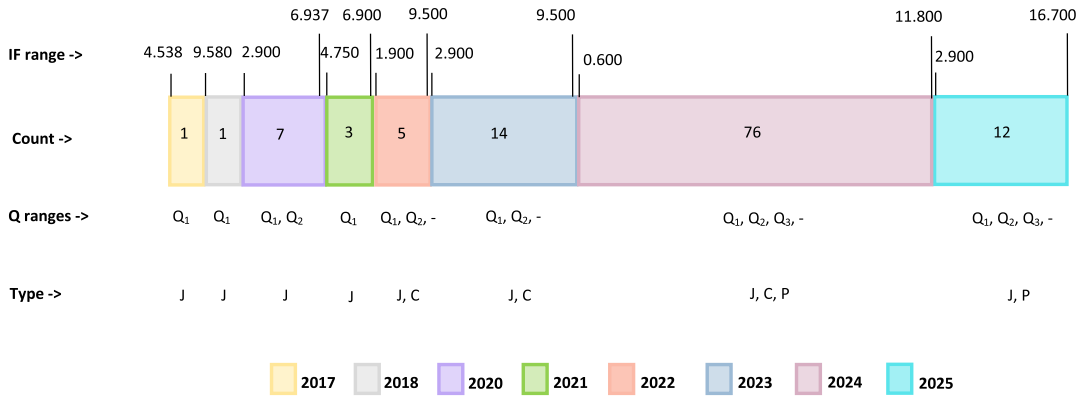


Figure 1. Overview of the studied journal (J), Conference (c), and preprints (P) articles based on impact factor (IF) range, count of articles, and Q ranges for each year.

chemical structures and human language (e.g., synthesis procedures), has further expanded NLP's potential, with LLMs like ChatGPT[54] being explored as tools for chemical reasoning and DD.[46]

SMILES, a chemical string used in drug discovery, represents molecular structures through ASCII characters, describing atoms, connectivity, bond order, and chirality.[55] This notation system is crucial for computational applications, including NLP.

The convergence of NLP and chemistry accelerates drug discovery by enabling efficient analysis of large chemical data sets, predicting molecular behaviors, and bridging chemical and natural language modeling.[50] Thus, NLP enhances our understanding of the "language of chemistry" and shapes pharmaceutical research and development.[46]

Drug discovery involves various branches aimed at improving disease treatment. A key focus is predicting DTI, which involves interactions with proteins or viruses.[56] However, DTI is not the only area of interest; DDI are also crucial for effective treatment.[57] DDI is closely related to ADR, where DDI examines drug interactions, and ADR focuses on harmful side effects, either from a single drug or multiple drugs in combination.[58]

The significance of computational drug discovery has led to surveys exploring the role of AI in this field. This survey reviews over 100 papers published from 2020 to 2025, focusing on DDI, DTI, and ADR extraction. Figure 1 highlights the progress in drug repurposing methods, while Table 3 lists the studied papers. Although existing reviews address these topics, they often focus on only one aspect—e.g., DDI[59−62] or DTI.[63] Some reviews like[64] critique the lack of structured categorization, while others limit themselves to machine learning, especially deep learning.[62,65] Notably, there are few studies on ADR.

In contrast, our survey covers the interconnected topics of DTI, ADR, and DDI. It includes a broader range of models, such as those using NLP techniques for biomedical text data sets and non-NLP models for chemical data sets. We categorize NLP-based models as "Record-based" and non-NLP models as "Structured-based", providing a more comprehensive view. This approach bridges gaps left by previous reviews, offering a thorough analysis of current research in DDI, DTI, and ADR extraction.

This survey systematically examines the use of AI methods in drug repurposing, excluding DTIs involving viruses, as guided by ref 66.

In this survey, we provide a background in Section 2 that defines ADR, DDI, and DTI. The next section is dedicated to databases and data sets in Section 3 used in the studied models. Next, we will have a review of some models in Section 4 and summing up all of the methods used in their structure, then, evaluation of models in Section 5. Furthermore, the challenges in Section 6 in the drug discovery field are summarized, and finally, we will have an outlook section in Section 7 to discuss the drug discovery opportunities and sum up the survey.

2. BACKGROUND

This section outlines the foundational concepts of DTI, DDI, and ADR. It also introduces key methodologies such as AI, machine learning, deep Learning, graph learning, and NLP, which are widely employed in drug repurposing research. The advantages, limitations, and future directions of each approach are briefly discussed. Expert readers may choose to skip this section and proceed directly to the subsequent parts of the paper. Additionally, we introduce AI models in the Supporting Information for interested readers.

2.1. Adverse Drug Reaction (ADR). Adverse drug events (ADEs) refer to unintended and undesired reactions to medications and have become more prevalent with the rise in the variety of drugs available worldwide since the 1950s.[67−69] Despite various safety and healthcare systems in place, medical professionals often report adverse events late, underscoring the need for systems that can extract necessary data from biomedical research and articles.[70] For instance, Thalidomide, used in the 1950s and 1960s to treat morning sickness during pregnancy, caused abnormalities in newborns, affecting over 10,000 cases across 17 countries.[71−74] Another example is the diet pill Fenfluramine, which was popular in the early 1980s but had adverse effects on the heart valves and rhythm of hundreds of thousands of people in the United States.[74−76]

ADRs, a subset of ADEs, are harmful side effects that occur when medications are taken at the recommended dose, including lack of efficacy, that occur during their use for prevention, diagnosis, or treatment of illnesses, or for modifying physiological functions.[77−81] Despite advancements in medicine over the past three decades, the rate of ADRs has remained consistent.[80] As laboratory samples are small, and initial clinical drug trials have limited observation time and scope, many potential adverse reactions go undetected.[74] These reactions contribute to 5−10% of hospital admissions globally and are a leading cause of death worldwide. In Australia, ADRs account for 2−4% of hospital admissions and are a significant source of patient illness.[80] Also, 4.2−30% of hospital admissions in the USA and Canada, as well as 2.5−10.6% of hospital admissions in

https://doi.org/10.1021/acsomega.5c04997
ACS Omega 2025, 10, 61048−61075

Table 1. Datasets Used in the Studied Articles

type	data set	models
ADR	FDA drug label Link	173
	ADE-corpus-v2[174] Link	142
DDI	DDI-Extraction 2013	175,176,177,178,177,179,96,151,70,95,180
	TAC 2018	95
	TAC 2019	101,95
	Deng[181]	182,181,183
	PK DDI[184]	101,180
	CHEMDNER[185]	96
	ChChMiner Link	186
	ZhangDDI Link	186
	OGB-biokg[187]	188
	KEGG-drug[189]	188
DTI	Luo et al.[190]	192,193,194,195,196,190,197
	BIOSNAP[198]	199,148,200,201,202,203,204,205,206,207,208
	DAVIS	209,210,199,211,212,200,213,214,215,216,217,218,219,220,196,221,206,163,208,222,223
	KIBA	210,224,212,213,215,216,217,218,219,211,220,221,196,163,222
	DUD-E[225,226]	227
	Human	148,201,203,214,228,204,202,205,207,227
	DTPKS[229]	230
	Data setKS	230
	Data setKSE	230
	DRKG[231]	230
	LINCS L1000[232]	230
	OncologyScreen[233]	234
	C. elegans[235]	214,205,228,208
	Wang[236]	196
	Yamanishi[237,238]	118,194
	Zheng[239]	193
	PrimeKG (DS1, DS2, DS3)[240]	241
	HIV Drug Interaction[242]	243
	Qmugs[244]	212
	CASF-2016[245]	212
	DUD-E[225,226]	212
	DrugProt[246]	247
	248	224
		210
	249	249

Europe, are caused by ADRs. In the US and Australia, the estimated costs of ADEs and ADR management are annually 30.1−130B US dollars and 4.83−9.00B Australian dollars per year, respectively.[82−84] In 2021 alone, China reported 597,000 new and severe ADEs, 11% of which were serious, posing a significant threat to patients' lives.[74]

2.2. Drug−Drug Interaction (DDI). In the US, DDIs result in approximately 74,000 emergency room visits and 195,000 hospitalizations annually.[85,86] A 2007 study analyzing 23 global clinical studies found that DDIs accounted for 0.054% of emergency room visits, 0.57% of hospital admissions, and 0.12% of rehospitalizations.[86,87] Between 1988 and 1994, 39.1% of Americans took prescription medications; this figure increased to 47.5% between 2007 and 2010.[86,88] DDIs affect 20−40% of patients and are a frequent cause of ADRs, particularly in the elderly due to polytherapy.[89−92] Polytherapy complicates treatment and increases the risk of clinically significant drug interactions, which may lead to ADRs and alter clinical efficacy.[86,93,94]

DDIs are classified into two main categories:[92,95]

- Pharmacokinetic (PK): Involves absorption, distribution, metabolism, and excretion, potentially causing treatment failure or toxicity.
- Pharmacodynamic (PD): Includes effects on receptor function, interference with biological or physiological control processes, and additive or opposing pharmacological effects.

DDI extraction identifies and classifies drug interactions within a given sentence.[95,96] Text mining plays a vital role in medical science and healthcare systems by extracting valuable information from unstructured text.[97−101] However, traditional methods like pattern-based and feature-based models are limited in capturing textual context and are time-consuming.[95]

Although DDIs have received less attention than gene−gene or protein−protein interactions,[70] events like SemEval 2013-Task 9 have increased interest in the field.[70,102,103] Machine learning approaches show promise, but deep learning techniques such as CNNs and RNNs provide greater flexibility and performance for large NLP data sets.[70,104−107] By utilizing these methods, studies aim to construct databases of known DDIs by extracting novel findings from scientific articles.[70]

https://doi.org/10.1021/acsomega.5c04997
ACS Omega 2025, 10, 61048−61075

Table 2. Overview of SE Datasets

type	models	databases/web sites
ADR	250,[251,252],142,74,253,252,254	MIMIC-III[256]
		Twitter
		Health Web sites and social media
		MedDRA Link
		PubMed[257]
		ChEMBL[258]
		China Pharmaceutical Information Query Platform Link
		Commonwealth Bank Health Society (CBHS) Link
		SIDER[259]
		DrugBank[143] Link
DDI	260,179,261,262,263,264,265,186,266,267,268,269,175,270,271,180	DrugBank[143]
		TWOSIDES[272]
		PubMed[257]
		ChemProt[273]
		DDIMDL Link
		DAS-DDI Link
		KEGG[274]
		Offside[272]
		STITCH[275]
		PubChem[276]
		14 data sets from[277]
DTI	278,279,234,247,192,280,220,204,281,282,216,203,214,201,283,200,212,224,148,284,285,286,287,288,206,207,211	DrugBank
		UniProt[140]
		MalaCard[289]
		BindingDB[146]
		HPRD[290]
		CTD[291]
		SIDER[259]
		KEGG[274]
		sc-PDB[292]
		Protein Data Bank (PDB) Link
		Binding MOAD database[293]
		PDBbind database[294]
		2P2I[295]
		GDSC[296]
		COSMIC[297]
		PubChem[276]
		pdbBIND[294]
		GPCR[298]
		ChEMBL[258]
		PharmGKB[299]

2.3. Drug−Target Interaction (DTI). The term "drug repurposing" was introduced in 2004 to address the limitations of traditional drug discovery methods,[108] though its origins date back to the 1950s.[109] This approach focuses on identifying new uses for existing and approved drugs.[110] Drug repurposing offers significant potential, especially given the high costs and long timelines of developing new drugs.[110,111] Research in this field has grown substantially due to advances in biological sciences and interdisciplinary fields such as bioinformatics, with increasing interest since 2011.[112] For example, "Sildenafil citrate" initially developed in 1980 for treating chest pain due to restricted blood flow,[113] was ineffective in clinical trials, halting its market introduction. However, reports of a side effect—prolonged erections led Pfizer to re-evaluate the drug in 1998. After trials with more than 3700 male volunteers showed positive results, Sildenafil was rebranded as "Viagra" to treat sexual disorders, becoming a best-selling drug with annual sales exceeding 1.5 billion dollars.[114,115]

One of the most important steps in drug discovery is accurately predicting DTIs.[116,117] So, prediction using computational techniques is vital for drug repurposing, as it is more cost-effective and time-efficient than biochemical experimental methods.[118] Drug developers screen compounds for interactions with targets showing desirable biological activities.[117] DTI can identify new drugs for a specific target or uncover new targets and applications for known drugs.[119] However, identifying DTIs through experiments typically takes 2−3 years and incurs significant costs.[116,120−122] Several computational techniques are created for DTI prediction, and a crucial problem is how to predict it accurately and quickly.[119] Consequently, as data on drugs, targets, and interactions continues to grow, various computational methods have been

https://doi.org/10.1021/acsomega.5c04997
ACS Omega 2025, 10, 61048−61075

developed to predict potential DTIs, enhancing the efficiency of drug discovery.[121]

Traditional methods for predicting DTIs are primarily categorized into structure-based and ligand-based approaches. Structure-based methods often utilize molecular docking tools to identify new ligands for proteins with known 3D structures or to discover new protein targets for existing drugs. In contrast, ligand-based methods typically employ pharmacophore searches and similarity assessments based on 3D shapes, substructures, and physicochemical properties.[119,123−126]

Public databases enhance the effectiveness of computational strategies for DTI prediction.[127] Recently, machine learning methods have gained popularity for identifying potential drug targets, assuming similar drugs share similar targets.[128,129] These methods often utilize kernel-based approaches, which map various drug−drug and target−target similarity matrices to DTI labels.[128,129] A new trend is the use of graph-based methods, which improve upon kernel-based methods by better representing the interactive relationships between drugs and targets through vertices and edges and extracting topological features from DTI networks for DTI predictions.[118,130] Also, deep learning methods have recently been applied extensively to DTI prediction and have shown notable performance gains.[64,117,121]

3. DATABASES AND DATA SETS

In the field of computational drug discovery, several common data sets are frequently used. These data sets typically include designated training and test sets. However, some researchers opt to use databases to create their custom data sets, which we will refer to as "SE" in this context.

Tables 1 and 2 present the common data sets and databases employed by researchers, respectively.

3.1. Databases. *3.1.1. UniProt.* The UniProt database[1] is a pivotal resource for protein sequence and functional information, boasting a capacity of 80 million sequences, which has doubled recently. Key features include the distinction between UniProtKB/Swiss-Prot (manually curated)[131,132] and UniProtKB/TrEMBL (unreviewed)[133,134] sections, as well as the introduction of a proteome identifier for unique organism assemblies.[135] Recent enhancements encompass a redesigned Web site focused on user experience, an annotation score to evaluate protein characterization, and automatic annotation systems like UniRule[136] and SAAS[137] to manage the entry of new sequences. With extensive cross-referencing to over 150 other databases and continuous effort toward high-quality curation based on literature, UniProt supports scientific discovery and offers valuable data for a broad spectrum of research fields.[138−140]

3.1.2. PubMed. PubMed is a free archive of biomedical and life sciences journal literature and is considered a reference for biomedical publications.[141] Some publications require licensed access but still provide a free version of the corresponding abstract. Therefore, PubMed was used as the main data source for literature monitoring to build the system.[142]

3.1.3. DrugBank. The DrugBank[2], introduced in 2006, served as a pioneering resource aimed at compiling comprehensive drug-related data, including information on FDA-approved drugs, investigational drugs, and their interactions.[143] This initial release focused on providing a centralized platform for drug data that included various pharmacological properties, biochemical pathways, and DTIs, facilitating research and education in pharmacology and related fields.[144] DrugBank 6.0, launched in 2024, significantly increased its resources such as FDA-approved drugs, investigational drugs, and DDIs, which have risen from 2646 to 4,563, from 3394 to 6,231, and from 365,984 to over 1.4M, respectively; it also features detailed pathways for drug mechanisms and improved accessibility through a refined user interface.[144] DrugBank 5.0, 2018 version, reported a 300% increase in investigational drugs and nearly 600% rise in DDIs, alongside new data sets in pharmacometabolomics and pharmacotranscriptomics, aiming to enhance the quality of drug information and support pharmacology research and education.[143]

3.1.4. BindingDB. BindingDB[3] is a comprehensive and publicly accessible database that has expanded significantly since its inception in 2001, now housing approximately 2.9 million binding measurements related to small molecules and proteins, making it essential for research in drug discovery and cheminformatics.[145−147] Its advanced search capabilities, user-friendly interface, and support for both manual curation and user contributions enhance its role as a vital resource for medicinal and computational chemistry, facilitating the development of new therapeutics through improved access to binding affinity data.[146]

3.1.5. DDInter 2.0. DDInter 2.0[4], designed to enhance clinical decision-making and patient safety, expands its predecessor's scope by covering 2310 drugs and 302,516 DDI records. It also includes 857 Drug−Food Interactions (DFIs), 8359 Drug−Disease Interactions (DDSIs), and 6033 Therapeutic Duplications, along with 8398 high-quality mechanism descriptions and management recommendations.[148]

The redesigned interface offers advanced filtering for mechanism- and risk-based screening, improving accessibility for healthcare providers. Integration with ADMET 3.0 enhances predictive capabilities, including Cytochrome P450 interactions.[149,150] DDInter 2.0 provides a comprehensive, freely accessible platform for managing drug interactions and improving patient outcomes.[148]

3.2. Data Sets. This section offers an overview of the data sets employed in the papers being reviewed. Certain studies may utilize either the entirety or a segment of the data sets listed below. We present the general details of each data set here. The Supporting Information includes specifics on how each reviewed paper managed its respective data set.

3.2.1. DDI-Extraction 2013. One of the well-known data sets used in DDI tasks is the DDI-Extraction 2013 corpus, an updated version of the Extraction 2011 data set provided by SemEval DDI 2013.[95,151] This data set is a collection of information on drug interactions and substances, derived from two different sources: DrugBank and PubMed article abstracts that are stored in MEDLINE.[70] The DrugBank subset comprises 730 documents, while the MEDLINE subset contains 175 abstracts for DDI extraction, further divided into training and testing sets of 714 and 191 documents, respectively.[95] Additionally, it includes 18,502 manually annotated drug entities and 5028 DDIs.[70]

3.2.2. TAC 2018 and 2019. TAC2018[5] and TAC2019[6] are curated data sets by the US Food and Drug Administration (FDA) and the National Library of Medicine (NLM),[152] designed to facilitate DDI research. TAC 2018 includes 325 Structured Product Labels (SPLs), divided into 22 training and 57 testing labels, categorizing DDIs into PK, PD, and Unspecified (U).[152,153] It also incorporates diverse data, such as articles and social media content.[95,154] TAC 2019 contains 400 SPLs in XML format, with 211 training and 81 testing labels, adding a "none" category for no DDI.[101,153] It includes four

https://doi.org/10.1021/acsomega.5c04997
ACS Omega 2025, 10, 61048−61075

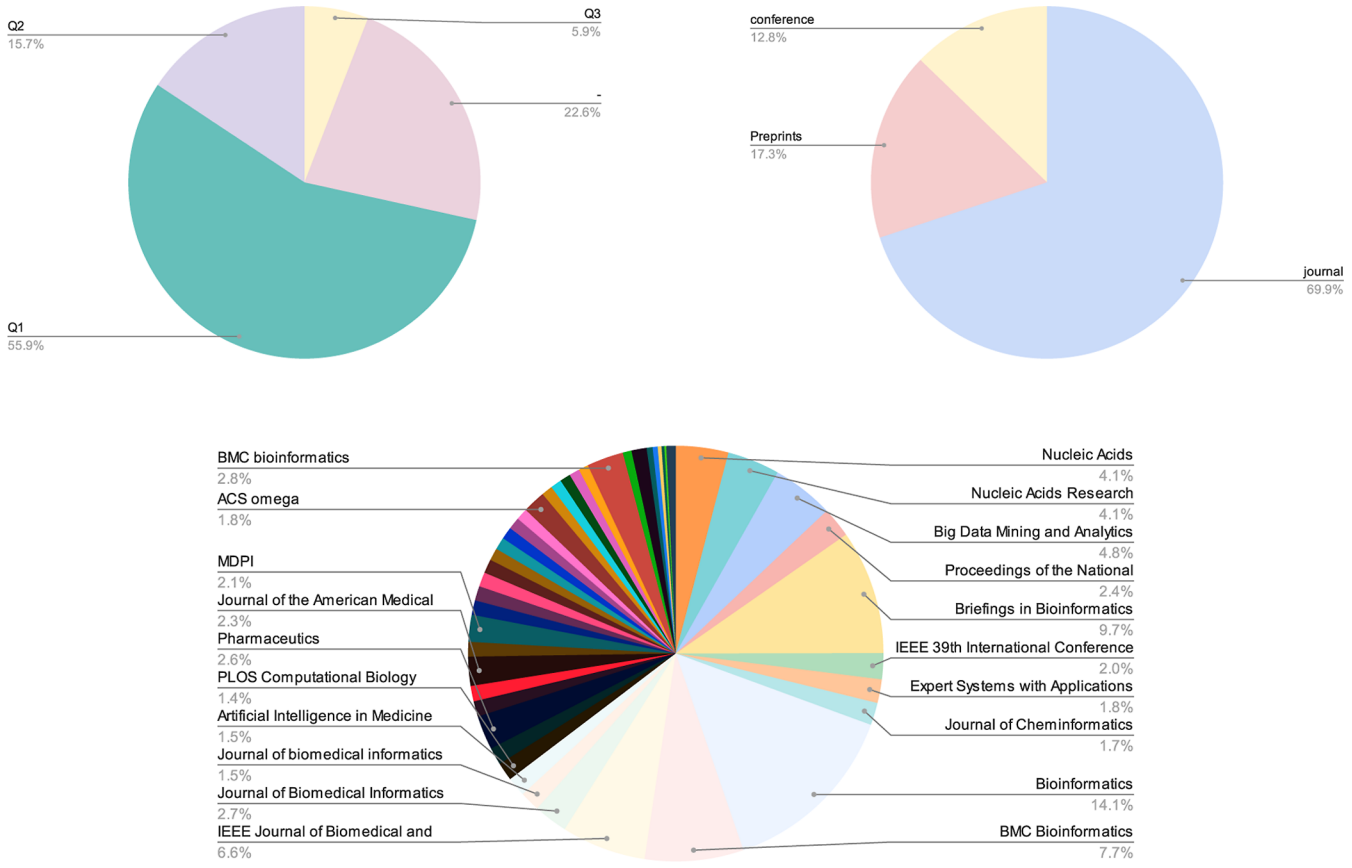


Figure 2. Percentage of articles by source name, ranks, and types.

training data sets, with 22 FDA labels fully annotated per TAC 2018 guidelines, plus reannotated labels for broader DDI classification.[95,101]

3.2.3. ADE-Corpus-V2. ADE-Corpus-V2 contains more than 20,000 sentences extracted from PubMed and prelabeled for drug-AE causality classification.[142] The ADE-Corpus-V2 is a specialized data set designed for pharmacovigilance research, focusing on ADE identification and extraction. Derived from medical case reports and summaries, the meticulously annotated corpus contains rich information on drug usage and associated ADEs. The data set comprises 4271 documents, covering 5063 drugs and 6821 ADE instances. ADE-Corpus-V2 is a valuable resource for developing and evaluating NLP models in drug safety research.[155]

3.2.4. CMED. Accurate medication history is essential for healthcare, but critical details often appear only in unstructured clinical notes, making analysis challenging.[156] Prior research classifies medication changes without full contextual labeling, a gap addressed by the Contextualized Medication Event Data set (CMED).[157]

CMED enhances medication timeline generation and reconciliation by annotating 9013 medication mentions from 500 clinical notes.[157−159] It categorizes events by Action, Temporality, Certainty, and Actor.[157,160] Developed using clinical notes from the 2014 i2b2/UTHealth NLP shared task, CMED's 2021 release included annotation guidelines and baseline assessments.[157,161]

3.2.5. DAVIS. The DAVIS data set provides experimental results on how kinase proteins interact with their inhibitors, including measurements called dissociation constants (K_d values). This data set includes information on the strength of interactions for 442 target proteins and 68 chemical compounds, with K_d values ranging from 5.0 to 10.8.[162,163]

The Davis data set contains pairs of kinase proteins and their respective inhibitors with experimentally determined dissociation constant (K_d) values. The data set comprises 442 proteins and 68 ligands, consisting of a total of 30,056 samples. The K_d values were transformed into pK_d values using the equation: pK_d = $-\log_{10}(K_d/1 \times 10^9)$.[164]

3.2.6. KIBA. The KIBA data set assesses how well drugs bind to their targets using a scoring system called the KIBA score, which is based on the analysis of related activity metrics. The KIBA data set features 467 target proteins and a large collection of 52,498 small drug molecules. After filtering, it includes 2111 distinct drug molecules and 229 distinct target proteins.[163,165] The KIBA data set comprises scores from different sources (K_i, K_d, and IC50) combined using the KIBA approach. Initially, the data set contained 467 proteins and 52,498 ligands, but was filtered to keep only drugs and targets with at least 10 samples. The final data set has 229 unique proteins and 2111 unique ligands, with a total of 118,254 samples.[165]

3.2.7. LINCS. NCBI GEO provided LINCS data for a landmark gene set of 978 genes (accessions GSE92742 and GSE70138), which included expression data from 83 cell types treated with 21,280 chemicals at varying doses and periods, 19,692 of which were matched to SMILES codes.[166,167] The expression data were analyzed using linear regression versus compound doses, resulting in pooled Z scores for all cell types and treatment periods that were subsequently discretized. For predictive modeling, each compound's transcriptional profile was coupled with a less comparable decoy profile, ensuring that

https://doi.org/10.1021/acsomega.5c04997
ACS Omega 2025, 10, 61048−61075

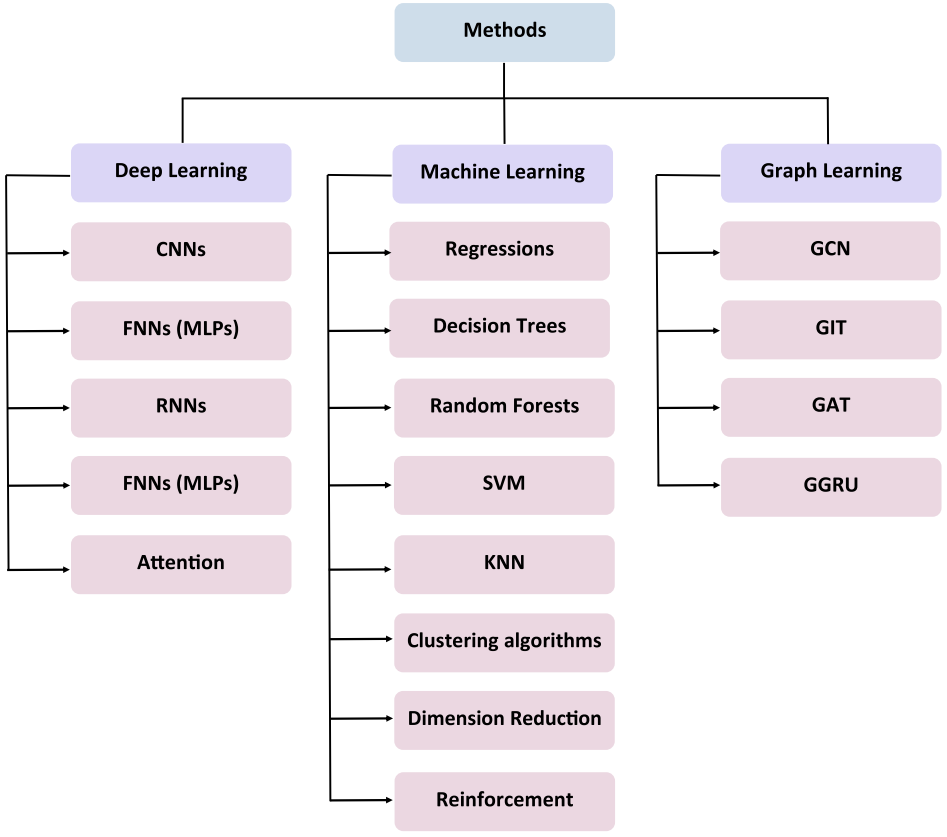


Figure 3. Overview of the general models for each method selected from papers.

the training and testing data sets did not contain any overlapping compounds.[168]

3.2.8. Human. The 2001 draft human genome sequence provided a foundational human data set, covering about 94% of the genome.[169] It is estimated that 30,000−40,000 protein-coding genes, with alternative splicing greatly increasing protein diversity.[170] The data set revealed that nearly half of the genome consists of transposable elements, showed segmental duplications near centromeres and telomeres, and suggested bacterial horizontal gene transfer in evolution. It also highlighted higher mutation rates in males and distinct recombination patterns, and identified over 1.4 million SNPs.[171] Using a hierarchical shotgun sequencing approach, the project ensured open data access, accelerating genomic research and future advancements.[172]

Some self-extracted (SE) data sets refer to collected data sets from different databases or Web sites and Table 2 summarizes them.

4. METHODS

This systematic literature review was conducted, adhering to the established guidelines outlined in ref 66. To gather relevant conference and journal papers, our strategy involved screening the Google Scholar search engine for English-language publications, primarily focusing on papers published from 2020 to 2025, with a strong emphasis on those in 2024. These articles are extracted from different journals, conferences, and preprint versions of arXiv[7] and BioRxive[8] (statistics about the articles are shown in Figure 1 and Figure 2), etc. The keywords utilized in the search included: "Drug−Drug Interactions", "Drug−Target Interactions", "DDI extraction", "DTI prediction", "Deep neural DDIs", "Machine learning DDIs", "Machine learning DTIs", "Adverse drug reactions", "Adverse drug events", and "ADRs extraction".

As previously discussed, the methods were categorized based on the employed models (Figure 3). Four categories were considered: Deep Learning models (i.e., CNN-based and RNN-based), Machine Learning models (regression and RF (Random Forest)), graph learning models (GCN) and hybrid models, which integrated deep, machine and graph approaches.

To illustrate the functionality of various approaches, a concise overview of several existing DDIs, DTIs, and ADRs prediction models will be provided. These examples serve to elucidate the diverse methodologies employed within the field. For a comprehensive analysis, including detailed information on over 100 relevant publications, refer to the Table 3. This table offers a comprehensive compilation of the literature, enabling a thorough understanding of the current state-of-the-art in computational drug discovery.

4.1. A Few Examples for ADR Extraction Model. First, it is important to clarify that this review specifically focuses on ADE-related publications that address ADRs. Consequently, to respect the authors' original terminology and perspective, we may use "ADE" interchangeably with "ADR" in certain instances. This will ensure consistency with the source material while maintaining clarity regarding the context of ADR prediction and analysis.

ADE extraction from social media involves three components: a RoBERTa-based classifier, a span extractor identifying ADE mentions in tweets, and a normalizer mapping them to MedDRA concepts.[251] FDA drug label ADE detection employs CRF, biLSTM + CRF, and BERT models, with ensembles improving accuracy. These models predict OSE ADE, non-OSE ADE, and Not ADE Candidate categories.[173] Predicate-ADR

https://doi.org/10.1021/acsomega.5c04997
ACS Omega 2025, 10, 61048−61075

Table 3. All Methods

model	output	approach	year	NLP	E2E	encoder	decoder	loss func	ratio pos2neg
250	ADR(c)	deep L	2020	√	√	for embedding one of the Glove, w2v, FastText, and BERT, BiLSTM	attention, GRU, MLP	cross-entropy	1:2
173	ADR(c)	deep L	2020	√	×	biLSTM, BioBERT	NB	-	-
251	ADR(c)	deep L	2021	√	√	BERT, GRU, MLP	RoBERTa	cross-entropy	-
142	ADR(c)	deep L	2022	√	×	Aho-corasick, BERTs	BERTs, fuzzy matching	-	-
74	ADR (c)	deep L	2023	√	×	ALBERT, BiLSTM	CRF	-	-
253	ADR(c)	hybrid	2024	×	√	w2v, n2v	ML classifiers (logistic regression, RF, KNN, SVM, XGBoost)	-	-
254	ADR(b)	hybrid	2024	×	×	DistMult,[302] HolE,[303] CNN	FC (one of: LR, RF, KNN, DT, NB, CNN, GBDT)	cross-entropy	12435/119,131 = 0.14
304	ADR(b)	-	2024	×	√	-	-	-	1:7.8
255	ADR(b)	deep L	2025	×	×	WAEs	FNN	Huber	-
260	DDI(c)	deep L	2018	√	×	PCA, DNN	SSP	cross-entropy	-
95	DDI(c)	deep L	2020	√	×	BioBERT	BioBERT, BiLSTM	-	1:5.9,1:4.9
183	DDI(n)	deep L	2020	×	×	Jaccard similarity, CNN	MLP	cross-entropy	-
181	DDI (b)	deep L	2020	√	×	Jaccard similarity, VGG16	MLP	cross-entropy	-
70	DDI(c)	deep L	2022	√	×	w2v, attention, GRU	RNN	cross-entropy	-
151	DDI (c)	deep L	2022	√	×	BiGRU, CNN, Max pooling, attention pooling	MLP	cross-entropy	-
101	DDI(n)	deep L	2023	√	×	BERT	BERT	-	-
96	DDI(b)	deep L	2023	√	√	BioBERT, w2v, d2v, TF-IDF, N-gram	various Machine Learning classifiers: Logistic Regression, SVM.	cross-entropy	-
182	DDI(n)	deep L	2023	×	×	Jaccard similarity, CNN	MLP	focal	-
261	DDI(b)	deep L	2023	×	√	GNN, ESPF and k-mer, attention	CASTER	cross-entropy	1:1
179	DDI(c)	deep L	2024	√	×	BERT	SVM, KNN, RF	-	-
262	DDI(c)	deep L	2024	×	×	attention or GAT (for d) and autoencoder (for binding), PCA, GCN	MLP	cross-entropy	-
177	DDI(c)	hybrid	2024	√	×	GloVe, BiLSTM, Attention, GGNN,UGC, CNN	linear layer	focal	-
263	DDI(n)	hybrid	2024	√	×	GCN,GAT, Jaccard similarity	average pooling layer, sigmoid	cross-entropy	-
264	DDI(n)	deep L	2024	√	√	attention, Pykeen deep learning framework, TransE algorithm, convolutional module	MLP	cross-entropy	
265	DDI(n)	hybrid	2024	×	×	knowledge graph, GNN	MLP	cross-entropy	-
186	DDI(b)	hybrid	2024	×		Tanimoto coefficient, RDkit, attention, RDkit, GAT, GCN	attention	cross-entropy	-
266	DDI(n)	hybrid	2024	×	×	DICnet module, se GNN, attention	MLP	multitask loss function	-
188	DDI(b)	hybrid	2024	×	√	KG2ECapsule, CapsGNN, CNN	MLP	cross-entropy	1:1
178	DDI(c)	deep L	2024	√	√	BioBERT, attention, LM, BIOES tagging, DNER module	MLP	cross-entropy	-
267	DDI(b)	deep L	2024	×	×	GAT, attention, GNN	sigmoid function	cross-entropy	-
175	DDI(c)	deep L	2024	√	×	Bert, BiLSTM, Scifive	MLP	cross-entropy	-
176	DDI(c)	deep L	2024	√	×	SubAGCN, attention, BiLSTM, CNN	Softmax layer	multifocal	-
177	DDI(c)	hybrid	2024	×	×	GLoVe, BiLSTM, attention, CDM, GCN	linear layer	Focal, margin	-
268	DDI(b)	hybrid	2024	×	×	GraphSAGE, Node2Vec, GNN, attention	MLP	cross-entropy	1:1
269	DDI(b)	hybrid	2025	×	×	attention, Mean shift algorithm, Conditional Domain Adversarial Network (CDAN), GAT	MLP	cross-entropy	1:1

https://doi.org/10.1021/acsomega.5c04997
ACS Omega 2025, 10, 61048−61075

Table 3. continued

model	output	approach	year	NLP	E2E	encoder	decoder	loss func	ratio pos2neg
270	DDI(b)	hybrid	2025	×	√	BRICS alg., GAT, HDN Blocks, attention	attention	cross-entropy	1:1
271	DDI(b)	hybrid	2025	×	×	attention, PCA, NNPS,[305] BERT	MLP	cross-entropy	-
180	DDI(b)	deep L	2025	√	×	LLMs	-	-	1:1
190	DTI(n)	machine L	2017	×	×	DCA	inductive Matrix completion	-	-
206	DTI(n)	deep L	2020	×	√	ChemBERTa (for chemical compounds), ProtBERT (for amino acid sequences)	MLP	SmoothL1	-
288	DTI(n)	deep L	2021	√ (inspired)	√	Ngram, CNN	MLP	-	-
118	DTI(b)	hybrid	2021	×	×	BiNE,[306] heuristic algorithm,path-based heterogeneous information	RF	cross-entropy	-
163	DTI(n)	deep L	2022	×	√	GRU (for p sequence), GNN (for d melcule)	MLP	MSE	-
164	DTI(n)	deep L	2022	×	×	one-hot encoding, CNN	MLP	MSE	-
227	DTI(b)	Graph	2022	√ (inspired)	√	GCNN with an attention pooling mechanism, attention	MLP	cross-entropy	3974731218 = 10.88
222	DTI(n)	hybrid	2023	×	×	GNNs(GAT, GCN, GIN, GINE, GMF)	MLP	MSE	-
278	DTI(n)	hybrid	2023	×	×	Cosine similarity, smith waterman algorithm, VGAE, GAE	RF	KL divergence, reconstruction error	-
279	DTI(n)	hybrid	2023	×	×	node2Vec, CNN	MLP	cross-entropy, triplet	-
207	DTI(n)	deep L	2023	×	×	GCN, LLM, attention	MLP	cross-entropy	BindingDB: 1.41(d),7.87(p); BioSNAP: 3.07(d), 6.34(p); Human: 1.23(d), 1.68 (p)
230	DTI(b)	hybrid	2023	×	√	attention, ConvE, GAT, GRU	MLP	cross-rnttopy	1:2
234	DTI(n)	hybrid	2023	×	×	n2v, attention, Dot product layer	MLP	cross-entropy, Lsynergy, LToxic	-
208	DTI(b)	deep L	2023	×	×	CNN, attention, k-gram	MLP	cross-entropy	0.84:1; 0.93:1; 0.98:1
223	DTA(n)	deep L	2023	×	×	RDKit, BiLSTM, attention, GIN	MLP	MSE	-
197	DTI(b)	deep L	2023	×	√	L-DLE, Jaccard similarity, attention	MLP	contrastive loss	1:1
196	DTI(c)	deep L	2024	×	×	CNN and n2v, attention	MLP	triplet	-
195	DTI(b)	graph L	2024	×	√	LM (ESM-2 and CHemBERT), GAT	MLP	cross-entropy	1:10
247	DTI(b)	deep L	2024	√	×	Pubtator, BERT	BERT	focal	-
205	DTI (n)	deep L	2024	×	√	BCM, CFM, CNN, attention	maximum pooling, MLP	cross-entropy	1:1
194	DTI (n)	machine L	2024	×	×	d similarity matrix, Similarity Selection	similarity matrix, random walk	-	1:1
193	DTI (n)	hybrid	2024	×	×	BIONIC,[307] attention, Jaccard similarity, GCN	MLP	cross-entropy	1:1, 1:5, 1:10 on all data sets
192	DTI (n)	hybrid	2024	×	×	Jaccard (for d), Tanimoto coefficient (for structures), Deepwalk, GraphSAGE	RF	cross-entropy	-
241	DTI (n)	deep L	2024	×	×	Attribute Masking, Average Pooling, attention, ComplEx embedding, GraphSAGE	MLP	focal, cross-entropy	-
243	DTI(n)	hybrid	2024	×	×	MLP, BPE, ChemBERT	MLP	cross-entropy	-
280	DTI(n)	deep L	2024	×	×	graph attention, latent information	MLP	cross-entropy, Kullback–Leibler divergence	-
220	DTI(n)	hybrid	2024	×	×	one-hot encoding, GIN, CNN, attention	MLP, max pooling	MSE, RWing[308]	-
204	DTI(n)	hybrid	2024	×	×	GCN, CNN, attention	MLP	cross-entropy	1:1
281	DTI(n)	machine L	2024	×	×	SVM	SVM	-	-
228	DTI(n)	hybrid	2024	×	√	PCA, Cosine similarity, GAT, attention	MLP	MSE	1:1
221	DTI(n)	deep L	2024	×	×	CNN, BiLSTM	MLP	MSE	

https://doi.org/10.1021/acsomega.5c04997
ACS Omega 2025, 10, 61048−61075

Table 3. continued

model	output	approach	year	NLP	E2E	encoder	decoder	loss func	ratio pos2neg
282	DTI(n)	machine L	2024	√	×	UMAP, CHEMBERT, BiGRU	regression	MAE, MSE, RMSE, RMSLE	
219	DT(n)	hybrid	2024	×	×	GVP, GIN, attention	MLP	MSE	-
309	DTI(n)	hybrid	2024	×	×	InfoGraph, CPCProt, Tanimoto coefficient, GCN, attention	sigmoid	MSE	-
310	DTI(n)	hybrid	2024	×	×	w2v, attention, GCN, random walk	MLP	cross-entropy	
218	DTI(n)	hybrid	2024	×	√	GCN, Fusion module, CNN, attention	MLP	MSE	-
217	DTI(n)	hybrid	2024	×	√	Morgan fingerprint, CNN, GVP, GNN, attention	MLP	MSE	-
216	DTI(n)	deep L	2024	×	×	MACCS, CNN, attention	MLP	MSE	-
212	DTI(n)	deep L	2024	×	×	ESM-fold, CNN, attention	MLP	cross-entropy, MSE	-
215	DTI(n)	deep L	2024	×	×	GCN, cross-scale	MLP	contrastive loss, MSE	-
203	DTI(n)	hybrid	2024	×	×	CNN, GNN, attention	MLP	cross-entropy	0.725 (binding), 1.014 (BioSNAP), 1 (Human)
214	DTI(b)	hybrid	2024	×	×	VGNAE, GCN, Hierarchical Graph Pooling with Structure Learning (HGP-SP), Fusion module	MLP	cross-entropy	1:2.5
213	DTI(n)	deep L	2024	×	×	VAE, GatedCNN, attention	MLP	MSE	-
201	DTI(n)	deep L	2024	×	√	GCN, ACmix, attention, BiIN	MLP	cross-entropy	1:2
283	DTI(n)	hybrid	2024	×	×	TAPEtokenizer, BERT, HAKE, GAT, cosine similarity	NIMC	contrastive loss, cross-entropy, triplet	1:all
200	DTI(b)	deep L	2024	×	×	ChemBERTa, ProtBERT	MLP, GeLU	cross-entropy	-
247	DTI(b)	deep L	2024	×	×	BERTs	MLP	cross-entropy	
224	DTI(n)	hybrid	2024	×	×	GCN, Attention, PLM, ChemBERTa	cross-attention	-	-
191	DTI(b)	hybrid	2024	×	×	3-g embedding, GNN, attention	MLP	cross-entropy	
148	DTI(b)	hybrid	2024	×	×	SELFormer, Saprot, SVM, RF, attention	MLP	cross-entropy	1:1
199	DTI(n)	hybrid	2025	×	√	GCN, attention, ChemBERTa	MLP	cross-entropy	-
210	DTI(n)	deep L	2025	×	√	BiGRU, fusion module, CNN, attention	MLP	MSE	-
284	DTI(b)	hybrid	2025	×	×	MLPS algorithm, attention, GCN	MLP	cross-entropy	15,250/17,341 = 0.879
249	DTI(n)	deep L	2025	×	×	attention, GNN	MLP	cross-entropy	1:1
285	DTI(b)	hybrid	2025	×	×	Self-supervised Learning, HGCN, GCN, attention	attention fusion, GCN	BPR	-
286	DTI(b)	hybrid	2025	×	×	Random Walk, CNN, BiAN (BiLSTM with attention)	MLP	cross-entropy	1:5
287	DTI(b)	hybrid	2025	×	×	USE, GCN, CNN, RNN, RNN	MLP, Logistic regression	-	-
209	DTI(b)	hybrid	2025	×	×	CNN, attention, BFIM, SFM	MLP	-	-
211	DTI(b)	hybrid	2025	×	×	CharVec, CNN	MLP	MSE	1:1
202	DTI(b)	hybrid	2025	√	×	LLMs (MoLFormer (4 d), ProtT5 (4 p)), GNN	MLP	cross-entropy	16,463/19,998 = 0.968

https://doi.org/10.1021/acsomega.5c04997
ACS Omega 2025, 10, 61048−61075

Table 4a. Evaluation of Studied Models (Accuracy, *P*, *R*, *F*1, ROC-AUC, PR-AUC)

output	model	accuracy	*P*	*R*	*F*1	ROC-AUC	PR-AUC
ADR	250	-	micro: 0.9455	micro: 0.9429	micro: 0.9442	-	-
			macro: 0.9377	macro: 0.9407	macro: 0.9379		
	173	-	0.8046 (CRF)	0.7879 (CRF)	0.7961 (CRF)	-	-
			0.8263 (BiLSTM-CRF)	0.8327 (BiLSTM-CRF)	0.8295 (BiLSTM-CRF)		
			0.7689 (BioBERT)	0.8345 (BioBERT)	0.8004 (BioBERT)		
	251	-	0.432 (GloVe)	0.171 (GloVe)	0.245 (GloVe)	-	-
			0.571 (w2v)	0.182 (w2v)	0.276 (w2v)		
			0.741 (FastText)	0.192 (FastText)	0.304 (FastText)		
			0.785 (BERT)	0.200 (BERT)	0.319 (BERT)		
	142	-	max: 0.95 (PubmedBert)	max: 0.93 (UMLSBERT and BioBERT)	Max: 0.93 (BioBERT SciBERT and Bio-Clinical Bert)	-	-
	74	-	0.8988	92.55	91.19	-	-
	253	0.8332 (RF)	0.8100 (RF)	0.8709 (RF)	0.8391 (RF)	0.9079 (RF)	-
		0.8148 (KNN)	0.7591 (KNN)	0.9243 (KNN)	0.8332 (KNN)	0.8925 (KNN)	
		0.8380 (LR)	0.8202 (LR)	0.8663 (LR)	0.8423 (LR)	0.9049 (LR)	
		0.8428 (SVM)	0.8094 (SVM)	0.8976 (SVM)	0.8509 (SVM)	0.9115 (SVM)	
		0.8443 (XGB)	0.8390 (XGB)	0.8532 (XGB)	0.8454 (XGB)	0.9157 (XGB)	
	254	0.886	0.867	0.914	0.891	0.959	-
	255	0.9604	0.9604	0.852	0.8875	0.927	
	304	-	-	-	-	-	-
	260	mean: 0.87 (from graph)	macro: 0.85	macro: 0.84	macro: 0.83	-	-
			micro: 0.87 (from graph)	micro: 0.86 (from graph)	micro: 0.84 (from graph)		
	95	-	-	-	weighted: 0.9179, macro: 0.8332 (DDI2013)	-	-
					weighted: 0.6053, macro: 0.8023 (TAC 2018)		
	183	0.8871	0.8556	0.722	0.7496	0.998	0.9251
	181	0.8852, (task 1)	0.8471 (task 1)	0.7182 (task 1)	0.7585 (task 1)	0.9976 (task 1)	0.9208 (task 1)
		0.6415 (task 2)	0.5607 (task 2)	0.4319 (task 2)	0.4460 (task 2)	0.9799 (task 2)	0.6558 (task 2)
		0.4075 (task 3)	0.2408 (task 3)	0.1452 (task 3)	0.1590 (task 3)	0.9512 (task 3)	0.3635 (task 3)
	70	-	0.8381	0.8159	0.8268	-	-
	151	-	0.7619	0.7334	0.7474	-	-
	101	-	0.827	0.806	0.816	-	-
	96	-	0.856849 (LR)	0.794636 (LR)	0.813400 (LR)	-	-
			0.851789 (SVM)	0.772669 (SVM)	0.795577 (SVM)		
			0.808146 (DT)	0.676117 (DT)	0.713850 (DT)		
			0.845322 (RF)	0.709323 (RF)	0.743582 (RF)		
			0.830406 (NB)	0.672542 (NB)	0.711865 (NB)		
			0.797973 (KNN)	0.619668 (KNN)	0.665580 (KNN)		
	182	-	0.9237	0.9237	0.9237	0.9986	0.9627
	247	-	-	-	0.806	-	-
	179	0.788 (ChemProt) 0.815 (DDI 2013)	-	-	-	-	-
	262	0.9589	-	-	0.9173	0.9992	0.9847
	261	-	-	-	0.8921 (TWOSIDES)	0.9625 (TWOSIDES)	0.9653 (TWOSIDES)
					0.9461 (DrugBank)	0.9869 (DrugBank)	0.9868 (DrugBank)
	177	0.726	0.697	0.758	-	-	-
	263	0.9120 (DS1)	0.8402 (DS1)	0.8578 (DS1)	0.8394 (DS1)	0.9523 (DS1)	0.9231 (DS1)
		0.8839 (DS2)	0.7353 (DS2)	0.7085 (DS2)	0.7329 (DS2)	0.9259 (DS2)	0.8470 (DS2)
	264	0.5129 (DS1)	-	-	0.4698 (DS1)	0.9649 (DS1)	0.4648 (DS1)
		0.6696 (DS2)			0.6537 (DS2)	0.9886 (DS2)	0.7197 (DS2)
	265	0.9409 (T1),0.6805 (T2) ———-	0.9132 (T1)	0.8876 (T1)	0.8958 (T1)	0.9989 (T1)	0.9786 (T1)
		note:T 4 Task	0.6063 (T2)	0.5106 (T2)	0.5394 (T2)	0.9673 (T2)	0.7049 (T2)
	186	0.9847	-	-	0.9845	0.9974	-

https://doi.org/10.1021/acsomega.5c04997
ACS Omega 2025, 10, 61048−61075

Table 4a. continued

output	model	accuracy	*P*	*R*	*F*1	ROC-AUC	PR-AUC
		(DrugBank) ————- 0.9509 (ChChMin) ————- 0.9548 (ZhangDDI)			(DrugBank) ———— 0.9704 (ChChMiner) ————- 0.9085 (ZhangDDI)	(DrugBank) ———— 0.9983 (ChChMiner) ————- 0.9898 (ZhangDDI)	
	266	-	-	-	-	-	-
	188	0.9519 (DS1-OGB-biokg) 0.9027 (DS2-KEGG-drug)	0.9371 (OGB-biokg) 0.8863 (KEGG-drug)	0.9715 (OGB-biokg) 0.9237 (KEGG-drug)	0.9528 (OGB-biokg) 0.9046 (KEGG-drug)	0.9928 (OGB-biokg) 0.9573 (KEGG-drug)	0.9921 (OGB-biokg) 0.9428 (KEGG-drug)
	178	-	0.9816 (DNER) 0.8985 (RC)	0.9865 (DNER) 0.8483 (RC)	0.9840 (DNER) 0.8726 (RC)	-	-
	175	-	0.875 (DDI 2013) 0.823 (ChemProt)	0.841 (DDI 2013) 0.774 (ChemProt)	0.849 (DDI 2013) 0.798 (ChemProt) 0.721 (DDIs data set)	0.603 (DDIs data set)	-
	176		macro: 0.8502 micro: 0.8922	macro: 0.8282 micro: 0.8072	macro: 0.8391 micro: 0.8475	macro: 0.9913 micro: 0.9796	macro: 0.9096 micro: 0.7300
	177	-	0.7789	0.8247	0.8012	-	-
	268	0.99	0.99	0.99	0.99	-	-
	269	0.9725 (DrugBank) 0.9443 (TWOSIDES)	0.9611 (DrugBank) 0.8631 (TWOSIDES)	0.9788 (DrugBank) 0.9668 (TWOSIDES)	0.9717 (DrugBank) 0.9245 (TWOSIDES)	0.9927 (DrugBank) 0.9563 (TWOSIDES)	-
	270	0.7984 (DrugBank) 0.8943 (TWOSIDES	-	-	0.8948 (DrugBank) 0.9592 (TWOSIDES)	-	-
	271	-	0.909	0.97	0.938	-	-
DDI	180	max: 0.7358 (0 shot Claude) 0.9260 (fine-tune GPT-4o)	max: 0.9212 (0 shot Qwen2.5) 0.9230 (Fine-tune Gemma2)	max: 0.9982 (0 shot LLaMa-3.2) 0.9690 (Fine-tune Qwen2.5)	max: 0.6957 (0 shot LLaMa-3.3) 0.9260 (Fine-tune GPT-4o)	-	-
	190	-	-	-	-	0.923 (extracted from chart)	0.935
	206	-	-	0.863 (BioSNP) 0.889 (DAVIS) 0.928 (BindingDB)	-	0.914 (BioSNP) 0.942 (DAVIS) 0.926 (BindingDB)	0.901(BioSNP) 0.517 (DAVIS) 0.639 (BindingDB)
	288	0.873	0.876	0.882	0.868	0.961	0.985
	118	-	-	-	-	Mean 0.9915 (Yamanishi) 0.429 (Olayan RS)	Mean 0.9545 (Yamanishi) 0.881 (Olayan RS)
	163	-	-	-	-	-	-
	164	-	-	-	-	-	0.450 (DAVIS) 0.601 (KIBA)
	227	-	0.951 (HUMAN)	0.975 (HUMAN)	0.963 (HUMAN)	-	0.971 (DUD-E Data set) 0.991 (HUMAN)
	222	-	-	-	-	-	-
	278	-	-	-	-	0.9840 (DS1) 0.9485 (DS 2)	0.8247 (DS1) 0.7302 (DS2)
	279	0.942	0.931	0.95	0.94	0.994	0.993
	207	0.911 (Binding) 0.876 (BioSNAP)	-	0.920 (Binding) 0.876 (BioSNAP)	-	0.97 (Human; from chart) 0.965(Binding) 0.941 (BioSNAP)	0.97 (Human; from chart)— 0.955(Binding) 0.942(BioSNAP)
	230	-	-	-	0.951	0.9706	-
	234	-	-	-	-	-	-

https://doi.org/10.1021/acsomega.5c04997
ACS Omega 2025, 10, 61048−61075

Table 4a. continued

output	model	accuracy	*P*	*R*	*F*1	ROC-AUC	PR-AUC
	208	-	0.961 (Human), 0.959 (C. elegans) 0.895 (DAVIS)	-	0.987 (Human) 0.992 (C. elegans) 0.922 (DAVIS)	0.989 (Human) 0.994 (C. elegans) 0.492 (DAVIS)	-
	223	-	-	-	-	-	-
	197	-	-	-	0.9871	0.9895	-
	196	0.971	0.935	0.991	0.962	0.991	0.985
	195	-	-	-	-	0.951 (DrugBank) 0.939 (BindingDB) 0.974 (Yamanishi) 0.944 (Luo)	0.953 (DrugBank) 0.934 (BindingDB) 0.966 (Yamanishi) 0.948 (Luo)
	247	-	-	-	0.806	-	-
	220	0.787 (BioSNAP) 0.950 (Human) 0.975 (C.elegans)	0.828 (BioSNAP) 0.968 (Human) 0.992 (Celegans)	0.727 (BioSNAP) 0.926 (Human) 0.976 (C.elegans)	-	0.857 (BioSNAP) 0.992 (Human) 0.996 (C.elegans)	0.872 (BioSNAP) 0.991 (Human) 0.996 (C.elegans)
	241	-	-	-	-	S1 (0.953) S2 (0.968) S3 (0.968) S4 (0.821) note S 4 setting	S1 (0.630) S2 (0.719) S3 (0.719) S4 (0.249)
	193	0.9340 0.9413 0.9539 (DS1-Luo) 0.8984 0.8986 0.8886 (DS2- Zhang) Mean 0.9275 0.9358 0.9465 (DS2- Yamanishi) note all ratios in cells are 1:1, 1:5, 1:10	-	-	-	0.9787 0.9813 0.9749 (DS1-Luo) 0.9546 0.9339 0.9134 (DS2- Zhang) Mean 0.9333 0.9308 0.9006 (DS3- Yamanishi)	0.9701 0.9075 0.8431 (DS1-Luo) 0.9497 0.9265 0.9160 (DS2- Zhang) Mean 0.9291 0.9290 0.9022 (DS3- Yamanishi)
	192	-	-	-	-	0.9818 (Luo) 0.9666 (new DS)	0.9839 (Luo) 0.9685 (new DS)
	241	0.9867	0.9788	0.9949	0.9868	0.9986	0.9985
	243	0.729 (-Sim) 0.776 (-ChemBERTa)	0.703 (-Sim) 0.752 (-ChemBERTa)	0.621 (-Sim) 0.675 (-ChemBERTa)	0.642 (-Sim), 0.695 (-ChemBERTa)	-	-
	280	0.9821	-	-	0.9779	0.9949	-
	220	-	-	-	-	-	-
	204	0.896 (BindingDB) 0.836 (BioSNAP) 0.942(Human)	-	0.884(BindingDB) 0.816 (BioSNAP) 0.929 (Human)	0.900 (BindingDB) 0.840 (BioSNAP) 0.944 (Human)	0.960 (BindingDB) 0.909 (BioSNAP) 0.983 (Human)	0.947 (BindingDB) 0.907 (BioSNAP) 0.976 (Human)
	281	0.92	0.92	0.92	0.92	-	-
	228	-	0.9962 (C.elegans) 0.9762 (Human)	0.9699 (C.elegans) 0.9395 (Human)	0.9742 (C.elegans) 0.9443 (Human)	0.9964 (C.elegans) 0.9874 (Human)	0.9956 (C.elegans) 0.9876 (Human)
	221	-	-	-	-	-	-
	282	-	-	-	-	-	-
	219	-	-	-	-	-	-
	309	-	-	-	-	0.962	0.9605
	310	0.8836 (DTKG) 0.9325 (DAVIS) 0.9080 (KIBA)	0.8291 (DTKG) 0.8348 (DAVIS) 0.7820 (KIBA)	0.8211 (DTKG) 0.8805 (DAVIS) 0.7210 (KIBA)	0.8251 (DTKG) 0.8570 (DAVIS) 0.7503 (KIBA)	0.9429 (DTKG) 0.9660 (DAVIS) 0.9320 (KIBA)	0.8844 (DTKG) 0.8930 (DAVIS) 0.8370 (KIBA)

https://doi.org/10.1021/acsomega.5c04997
ACS Omega 2025, 10, 61048−61075

Table 4a. continued

output	model	accuracy	*P*	*R*	*F*1	ROC-AUC	PR-AUC
		00.8551 (DrugBank)	0.8137 (DrugBank)	0.8551 (DrugBank)	0.8339 (DrugBank)	0.9093 (DrugBank)	0.9136 (DrugBank)
	218	-	-	-	-	-	-
	217	-	-	-	-	-	-
	216	-	-	-	-	-	-
	212	-	-	-	-	-	-
	215	-	-	-	-	-	-
	203	0.907 (Binding)	-	0.908 (Binding)	-	0.965 (Binding)	0.957 (Binding)
		0.839 (BioSNAP)		0.819 (BioSNAP)		0.909 (BioSNAP)	0.909 (BioSNAP)
		0.945		0.944 (Human)		0.988 (Human)	0.984 (Human)
	214	-	0.981 (DrugBank)	0.954 (DrugBank)	-	0.989 (DrugBank)	0.992 (DrugBank)
				(DrugBank)		0.997(Human)	0.997(Human)
						0.996(C.elegans)	0.996(C.elegans)
			0.987 (Human)	0.976 (Human)		0.991(BindingDB)	0.993(BindingDB)
						−0.976 (GPCR)	0.983 (GPCR)
			0.984 (C. elegans)	0.981 (C. elegans)			
			0.976 (BindingDB)	0.943 (BindingDB)			
			0.964 (GPCR)	0.905 (GPCR)			
	213	-	-	-	-	-	0.6629 (DAVIS)
							0.7015 (KIBA)
	201	0.9130 (BindingDB)	-	-	0.9141 (BindingDB)	0.9655 (BindingDB)	0.9519 (BindingDB)
		0.8341 (BioSNAP)			0.8373 (BioSNAP)	0.9046 (BioSNAP)	0.9061(BioSNAP)
		0.8223 (DrugBank)			0.8249 (DrugBank)	0.8881 (DrugBank)	0.8884 (DrugBank)
		0.9407 (Human)			0.9398 (Human)	0.9794 (Human)	0.9753 (Human)
	283	0.936	0.1789	0.9168	0.2988	0.9751	0.631
	200	-	-	0.848 (BIOSNAP)	-	>0.914 (BIOSNAP)	0.914 (BIOSNAP)
				0.833 (DAVIS)		0.895 (DAVIS)	0.373 (DAVIS)
				0.888 (BindingDB)		0.912(BindingDB)	0.643 (BindingDB)
	247	-	-	-	0.806	-	-
DTI	224	0.942 (BindingDB E1)	High precision	Lower recall	High F1-score	-	-
		0.947 (BindingDB E2)	4 BindingDB E2 (no value)	4 GraphDTA:0.800 (BindingDB E2)	4 BindingDB E2(no value)		
		0.744 (DrugBank E3)					
	191	-	-	-	-	0.9805	0.976
	148	0.961 (BindingDB)	-	-	-	0.989 (BindingDB)	0.990 (BindingDB)
		0.889 (BioSNAP)				0.991 (Human)	0.989 (Human)
						0.951 (BioSNAP)	0.951 (BioSNAP)
	199	-	-	0.863 (BIOSNAP)	-	0.931 (BIOSNAP)	0.930 (BIOSNAP)
				0.837 (DAVIS)		0.934 (DAVIS)	0.462 (DAVIS)
	210	-	-	-	-		
	284	-	-	>0.799(GPCR)	-	0.860 (GPCR)	0.862 (GPCR)
				0.689 (DrugBank)		0.808 (DrugBank)	0.557 (DrugBank)
	249	-	-	-	-	0.9972	0.9976
	285	0.829	-	-	0.8	0.923	0.937
	286	-	-	-		0.9811	0.9868
	287	0.8258	0.9989	0.8258	0.9041	0.9143	-
	209	-	-	-		0.9746	0.9542
	202	-	-	0.866 (BioSNAP)	-	0.939 (BioSNAP)	0.941 (BioSNAP)
						0.898 (Human)	0.837 (Human)
DTI	211	-	-	-	-	-	0.7410 (Davis)
							0.8150 (KIBA)

models extract linguistic features from social media posts using feature pooling.[183] CNN-based models (CNN, HAN, FastText with Word2Vec) classify ADR comments, with preprocessing including case conversion, stop-word removal, and lemmatization.[300]

4.2. A Few Examples for DDI Extraction Models. DDI extraction has advanced with deep learning models like CNNs and RNNs, improving performance over traditional methods.[70] The SEV-DDI model, an LSTM-based approach, enhances accuracy using a dictionary-based method to predict DDI severity, aiding clinical decisions.[70]

https://doi.org/10.1021/acsomega.5c04997
ACS Omega 2025, 10, 61048−61075

Table 4b. Evaluation of Studied Models (Specificity, CI, R^2, MSE, MCC)

output	model	specificity	CI	R^2	MSE	MCC
ADR	250	-	-	-	-	-
	173	-	-	-	-	-
	251	-	-	-	-	between 1 and 2
	142	-	-	-	-	-
	74	-	-	-	-	-
	253	-	-	-	-	-
	254	-	-	-	-	-
	255	0.913	-	-	-	-
	304	-	-	-	-	-
	260	-	-	-	-	-
	95	-	-	-	-	-
	183	-	-	-	-	-
	181	-	-	-	-	-
	70	-	-	-	-	
	151	-	-	-	-	-
	101	-	-	-	-	-
	96	-	-	-	-	-
	182	-	-	-	-	-
	247	-	-	-	-	-
	179	-	-	-	-	-
	262	-	-	-	-	-
	261	-	-	-	-	-
	177	-	-	-	-	-
	263	-	-	-	-	-
	264	-	-	-	-	-
	265	-	-	-	-	-
	186	-	-	-	-	-
	266	-	-	-	-	-
	188	-	-	-	-	-
	178	-	-	-	-	-
	175	-	-	-	-	-
	176	-	-		-	-
	177	-	-	-		-
	268	-	-	-	-	-
	269	-	-	-	-	-
	270		-	-	-	-
	271	-	-	-	-	-
DDI	180	-	-	-	-	-
	190	-	-	-	-	-
	206	0.862 (BioSNP) 0.903 (DAVIS) 0.814 (BindingDB)	0.913 (BioSNP) 0.940 DAVIS) 0.927 (BindingDB)	0.467 (BioSNP) 0.201(DAVIS) 0.365 (BindingDB)	-	-
	288	-	-	-	-	-
	118		-	-	-	-
	163	-	0.75 (GCN) 0.70 (GAT) 0.83 (GIN)	-	0.75 (GCN) 0.77 (GAT) 0.57(GIN)	-
	164	-	0.876 (DAVIS) 0.836 (KIBA)	>0.637 (DAVIS) 0.614(KIBA)	0.276 (DAVIS) 0.226 (KIBA)	-
	227	-	-	-	-	-
	222	-	0.917 (DAVIS) 0.893 (KIBA)	0.722 (DAVIS) 0.784 (KIBA)	0.184 (DAVIS) 0.138 (KIBA)	-
	278	-	-	-	-	-
	279	-	-	-	-	-
	207	0.900 (Binding) 0.877 (BioSNAP)	-	-	-	-
	230	-	-	-	-	-
	234	-	-	-	-	-
	208	-	-	-	-	-
	223	0.907 (KIBA)	0.807 (KIBA)	0.124 (KIBA)	-	-

https://doi.org/10.1021/acsomega.5c04997
ACS Omega 2025, 10, 61048−61075

Table 4b. continued

output	model	specificity	CI	R^2	MSE	MCC
		0.914 (DAVIS)	0.744 (DAVIS)	0.196 (DAVIS)		
	197	-	-	-	-	-
	196	-	-	-	-	-
	195	-	-	-	-	-
	247	-	-	-	-	-
	220	-	-	-		-
	241	-	-	-	-	-
	193	-	-	-	-	-
	192	-	-	-	-	-
	241	-	-	-	-	-
	243	0.836 (DeepARV-Sim) 0.878 (DeepARV-ChemBERTa)	-	-	-	-
	280	-	-	-	-	-
	220	-	0.905 (DAVIS) 0.900 (KIBA)	-	0.220 (DAVIS) 0.134 (KIBA)	-
	204	0.913 (BindingDB) 0.857 (BioSNAP) 0.957 (Human)	-	-	-	-
	281	-	-	-	-	-
	228	-	-	-	-	-
	221	-	0.901 (DAVIS) 0.886 (KIBA)	-	0.215 (DAVIS) 0.147(KIBA)	-
	282	-	-	0.83 (ALL)	1.29 (ALL)	-
	219	-	0.895 (DAVIS) 0.886 (KIBA)		0.209 (DAVIS) 0.159 (KIBA)	-
	309	-	-	-	-	-
	310	-	-	-	-	-
	218	-	0.63 (DAVIS) 0.66 (KIBA) from image		0.66 (DAVIS) 0.56 (KIBA) from image	-
	217	-	0.903 (DAVIS) 0.892 (KIBA)	0.731 (DAVIS) 0.771 (KIBA)	0.177 (DAVIS) 0.14 (KIBA)	-
	216	-	0.890 (DAVIS) 0.860 (KIBA)	0.756 (DAVIS) 0.731 (KIBA)	0.199 (DAVIS) 0.184 (KIBA)	-
	212	-	0.646	-	0.547	-
	215	-	-	0.776(DAVIS) 0.808(KIBA)	0.166(DAVIS) 0.127(KIBA)	-
	203	0.906 (Binding) 0.858 (BioSNAP) 0.957 (Human)	-	-	-	-
	214	-	-	-	-	-
	213	-	0.8696 (DAVIS) 0.8221 (KIBA)	0.5713 (DAVIS) 0.6329 (KIBA)	0.3329 (DAVIS) 0.2536 (KIBA)	-
	201	-	-	-	-	0.8242 0.6694 0.6451 0.8800
	283	-	-	-	-	-
	200	0.844 (BIOSNAP) 0.802 (DAVIS) 0.793 (BindingDB)	-	-		-
	247	-	-	-	-	-
DTI	224	-	-	-	-	-
	191	-	-	-	-	-
	148	-	-	-	-	-
	199	0.857 (BIOSNAP) 0.884 (DAVIS)	-	-	-	-
	210	-	-	-	-	-
	284	-	-	-	-	-

https://doi.org/10.1021/acsomega.5c04997
ACS Omega 2025, 10, 61048−61075

Table 4b. continued

output	model	specificity	CI	R^2	MSE	MCC
	249	-	-	-	-	-
	285	-	-	-	-	0.687
	286	-	-	-	-	-
	287	-	-	-	-	-
	209	-	-	-	-	-
	202	0.857 (BioSNAP)	-	-	-	-
DTI	211	-	0.897 (Davis)	0.6240 (Davis)	0.242 (Davis)	
			0.890 (KIBA)	0.7070 (KIBA)	0.166 (KIBA)	

The RBioBERT-LSTM model combines RBioBERT with LSTM, surpassing RBioBERT alone in DDI classification.[95] BioBERT-directionalDDI, a BERT-based model, identifies PK DDIs and the interacting drug responsible for alterations.[301] The ALBERT-BiLSTM-CRF model improves medical entity recognition,[74] while LiSA screens biomedical literature for drug-adverse event relationships.[142] The Knowledge-based Att-BiLSTM model leverages UMLS and UIMA for entity detection, using Att-BiLSTM for relationship assignment.[250]

4.3. A Few Examples for DTI Prediction Models. A fine-tuned BERT model predicts DTIs using ChemBERTa and ProtBERT for encoding compounds and proteins, with an FCN decoder combining their outputs.[206] An end-to-end model employs GRU and GNN to extract drug-target features from SMILES sequences and molecule maps, merging them into a unified vector.[163] MPS2IT-DTI converts molecular and protein sequences into image-like matrices, applying CNNs to process them, differing from embedding-based models like DeepDTA.[164]

3DProtDTI integrates AlphaFold structure predictions with graph-based protein and ligand representations, using GNNs and FC layers for feature extraction.[222] A twin CNN model predicts DTIs from multimodal drug data, using a Siamese structure for learning chemical and biological features.[182] DTINet integrates drug−protein-disease data in a heterogeneous network, using random walk and dimensionality reduction.[190] FMCA-DTI applies a fragment-oriented multi-head cross-attention mechanism.[205]

Table 3 provides an overview of the studies reviewed, categorized by the models they employ. Note that all fully connected layers in the decoder are classified as MLP (Multilayer Perceptron). The table includes citations to each study, the type of output (ADR, DDI, or DTI), and the output format—categorized (c), numerical (n), or binary (b). The general approaches are classified as machine learning, deep learning, graph learning, or hybrid methods. The use of NLP tools is indicated in the "NLP" column, while the "E2E" column specifies whether the model operates end-to-end. The "Encoder" column describes the representation or embedding learning methods, and the "Decoder" column outlines the downstream prediction techniques. The subsequent column details the loss functions used. Due to the common imbalance in data sets, where negative samples often far outnumber positive ones, some studies select a smaller subset of negative samples. The final column, "Ratio p2n," reports the positive-to-negative sampling ratio applied in each study.

5. EVALUATIONS

Precision and recall assess systems' capability to identify relevant items, with precision focusing on correct retrievals and recall on relevant items found; their balance leads to the F-score. Metrics like accuracy, specificity, and MSE quantify overall performance, while the CI evaluates survival analysis. ROC-AUC visualizes performance, especially with imbalanced data sets, and the MCC measures binary classifiers' effectiveness, accommodating varying distributions. R^2 indicates how well a regression model explains variance in data. The Supporting Information contains the full description of evaluation metrics for interested readers.

By using these metrics, the models' evaluations are summarized in Table 4a and 4b. It should be noted that there were also other metrics used in these papers, such as Pearson[311] and Spearman.[312] However, only a few of them used these metrics; therefore, we removed them to save space in the table. Pearson was used in E-ADGCN[177] (0.232 on DAVIS and 0.215 on KIBA), and PocketDTA[217] (0.887 on DAVIS and 0.896 on KIBA), and Spearman was used in CSCo-DTA[215] with 0.231.

6. CHALLENGES

Deep learning models face several significant challenges in the pharmaceutical domain, particularly because of the need for large volumes of high-quality annotated data that are time-consuming and labor-intensive to obtain. Such challenges are compounded by model performance instability, influenced by noise and data imbalance, which complicates the understanding of complex sentences and diminishes the models' ability to capture critical meanings. Additionally, the interpretability of these models remains a concern, as the unpredictability of their outputs makes it difficult for pharmacists to identify potentially hazardous drug combinations, crucial for ensuring patient safety.[61] Specific issues such as sense ambiguity, matching errors, and relation classifier inaccuracies further hinder the accuracy of model predictions, affecting the reliability of data used in clinical settings.[250] Ambiguous concepts like "Anticoagulation", which can refer to either a drug or a condition, can result in classification errors; and matching errors may arise when systems incorrectly categorize terms due to overlapping descriptions, such as identifying "Cyanocobalamin 1000" as a medication while overlooking "1000 mcg/mL" as its strength.[250]

Beyond the technical hurdles, establishing sound business models for drug repurposing is essential, as challenges related to intellectual property rights and economic factors complicate drug approval and application for new uses.[313] Variations in patent laws across regions, such as those in the European Union and the United States, can impede the potential return on investment for new drugs.[314] Additionally, limited access to crucial data like clinical trial information, along with the complexities of data integration, presents further obstacles.[315] Moreover, the development of NER techniques is vital for identifying and classifying medical terms to enhance the detection of ADEs.[74] As the need for accurate DTI predictions grows, exploring mutual interactions between drugs and targets becomes increasingly important,[117] while advancements in NLP

https://doi.org/10.1021/acsomega.5c04997
ACS Omega 2025, 10, 61048−61075

highlight the necessity for effective detection of DDIs from the expanding body of biomedical literature.[101] Additionally, the heterogeneity and high dimensionality of biological and chemical data demand sophisticated computational methods, such as graph embedding, but these methods face difficulties in effectively integrating diverse data types. Furthermore, co-occurrence-based text mining can introduce false positives-identifying associations that are not biologically meaningful-necessitating careful filtering and validation.[316]

One related area in drug discovery is the study of drug−food interactions (DFIs). To improve public understanding of DFIs, it is important to create easy-to-understand educational resources, promote awareness of reliable scientific information, and ensure access to trustworthy sources. Drug discovery in DFIs includes developing accurate, user-friendly tools and apps that can provide personalized information on drug−herb interactions. Additionally, healthcare providers face challenges in effectively educating patients about potential food and medication interactions, which requires clear communication and guidance. Pharmacists play a crucial role in counseling patients, but ensuring consistent, accessible education and fostering positive attitudes toward understanding DFIs remain significant hurdles for advancing drug discovery and personalized medicine.[317] Another issue is the limited literature and manual curation of information, leading to data scarcity and inconsistency that can cause biases and incomplete data sets, which have been discussed before. Moreover, the complexity of units such as food composition makes it difficult to generate accurate, comprehensive feature representations for food constituents, preventing precise modeling. The limited availability of validated, gold-standard data sets for model training and evaluation also poses significant challenges for developing large, generalizable predictive models for DFIs.[316] Furthermore, healthcare professionals often have limited knowledge about drug−food interactions, which can lead to medication problems and health risks. Improving their education and awareness is essential to prevent such issues and ensure safer patient care.[318]

7. OUTLOOK

This survey presents the latest research studies on computational drug discovery, focusing particularly on DTI, DDI, and ADR. We provide an overview of the deep learning and LLMs utilized in this area of research. Additionally, we highlight relevant data sets and databases that are significant to these fields. Furthermore, we report on the specifications of recent methods proposed in this area and provide an evaluation and comparison of these methods for interested readers.

The high costs and extensive labor associated with annotating large data sets present challenges but also open opportunities for enhancing data utilization through various techniques. Implementing a separate concept reasoning module alongside machine learning models trained on gold-standard concepts can improve performance significantly.[250] Methods such as semisupervised and distant supervision modeling help address limited annotated data by using knowledge bases and bootstrap learning, although issues like incorrect labeling and data sparsity remain in biomedical applications. Advanced techniques like Generative Adversarial Networks (GANs)[319] and correlated information[320] can further mitigate these challenges and enhance performance in relation extraction tasks.[61]

Strategies like joint learning, ensemble learning, and dual learning can be employed to improve the stability and accuracy of relation extraction. Joint learning enables the simultaneous extraction of entities and relations, reducing error propagation, while ensemble learning combines predictions from multiple classifiers for greater accuracy. Additionally, dual learning can capitalize on task interrelations to boost performance. Feature enrichment using deep learning allows for the integration of rich data sources, including molecular information and social media content, leading to more robust and interpretable DDI extraction models.[61]

Table 5. Abbreviation Used in This Survey

Area under the precision−recall curve: AUPR	Recall: R
Area under the receiver operating characteristic: AUROC	Receiver operating characteristic: ROC
Bayesian Personalized Ranking: BPR	Simplified Molecular Input Line Entry System: SMILES
Biased Feature Interaction Module: BFIM	Sparse Feature Module: SFM
Context-Dependent Dense Module: CDM	Synergy Loss: Lsynergy
Decision Tree: DT	Toxic Loss: LToxic
Doc2Vec: d2v	Universal Sentence Encoder: USE
Gated Graph Neural Networks: GGNN	Wasserstein Autoencoders: WAEs
Gaussian Error Linear Unit	Feed-Forward Neural Networks
GeLU	FNNs
Geometric Vector Perceptron: GVP	Word2Vec: w2v
Gradient Boosting Decision Tree: GBDT	XGBoost: XGB
Graph attention network: GAT	bidirectional Intention network: BiIN
Graph neural network: GNN	binary: b
Hypergraph Convolutional Networks: HGCN	character: c
K-Nearest Neighbor: KNN	data set: DS
Logistic Regression: LR	end-to-end: E2E
Masked Language Modeling: MLM	low-dimensional latent embeddings: L-DLE
Naïve Bayes: NB	m Structured Product Labeling: SPL
Neural Inductive Matrix Completion	neural network-based predictor
NIMC	NNPS
Nod2Vec: n2v	numeric: n
P-glycoprotein: P-gp	pretrained large model: PLM
Precision: P	structural similarity profile: SSP
Precision−recall: PR	variational graph autoencoder: VGAE

8. ABBREVIATIONS

Abbreviations used in this survey are in the Table 5.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.5c04997.

The supplementary file includes an overview of some of the most-used machine learning, deep learning, and graph learning models. It also includes information on evaluation metrics used in the studied articles, as well as a table that outlines the train-validation-test divisions used in various proposed models (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Mohsen Hooshmand − *Department of Computer Science and Information Technology, Institute for Advanced Studies in Basic Sciences (IASBS), Zanjan 45137-66731, Iran;*

https://doi.org/10.1021/acsomega.5c04997
ACS Omega 2025, 10, 61048−61075

orcid.org/0000-0002-2934-2675;
Email: mohsen.hooshmand@iasbs.ac.ir

Author

Soheila Behrooznia − *Department of Computer Science and Information Technology, Institute for Advanced Studies in Basic Sciences (IASBS), Zanjan 45137-66731, Iran*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acsomega.5c04997

Author Contributions

M.H.: Supervision, Conceptualization, Writing—Review & Editing. S.B.: Conceptualization, Literature Search, Writing—Original Draft, Review & Editing, Visualization.

Notes

The authors declare no competing financial interest.

■ ADDITIONAL NOTES

[1]https://www.uniprot.org/.
[2]https://go.drugbank.com/.
[3]https://bindingdb.org/.
[4]https://ddinter2.scbdd.com/.
[5]https://tac.nist.gov/2018/.
[6]https://bionlp.nlm.nih.gov/tac2019druginteractions/.
[7]https://arxiv.org/.
[8]https://www.biorxiv.org/.

■ REFERENCES

(1) Hasselgren, C.; Oprea, T. I. Artificial intelligence for drug discovery: Are we there yet? *Annu. Rev. Pharmacol. Toxicol.* **2024**, *64*, 527−550.
(2) Zhang, S.; Liu, K.; Liu, Y.; Hu, X.; Gu, X. The role and application of bioinformatics techniques and tools in drug discovery. *Front. Pharmacol* **2025**, *16*, 1547131.
(3) Xia, X. Bioinformatics and drug discovery. *Current topics in medicinal chemistry* **2017**, *17*, 1709−1726.
(4) Catacutan, D. B.; Alexander, J.; Arnold, A.; Stokes, J. M. Machine learning in preclinical drug discovery. *Nat. Chem. Biol.* **2024**, *20*, 960−973.
(5) Croston, G. E. The utility of target-based discovery. *Expert Opinion on Drug Discovery* **2017**, *12*, 427−429.
(6) Zhang, J.; Li, C.; Lin, Y.; Shao, Y.; Li, S. Computational drug repositioning using collaborative filtering via multi-source fusion. *Expert Systems with Applications* **2017**, *84*, 281−289.
(7) Talevi, A. Tailored multi-target agents. Applications and design considerations. *Curr. Pharm. Des.* **2016**, *22*, 3164−3170.
(8) Berg, E. L. The future of phenotypic drug discovery. *Cell chemical biology* **2021**, *28*, 424−430.
(9) Taylor, D. The Pharmaceutical Industry and the Future of Drug Development. *Pharmaceuticals in the Environment* **2015**, 1−33.
(10) Torjesen, I. Drug development: the journey of a medicine from lab to shelf. *Pharm. J.* **2015**, *12*, 1−7.
(11) Wang, Y.; Deng, G.; Zeng, N.; Song, X.; Zhuang, Y. Drug-disease association prediction based on neighborhood information aggregation in neural networks. *IEEE Access* **2019**, *7*, 50581−50587.
(12) Hosseini, M.; Chen, W.; Xiao, D.; Wang, C. Computational molecular docking and virtual screening revealed promising SARS-CoV-2 drugs. *Precision clinical medicine* **2021**, *4*, 1−16.
(13) Aziz, M.; Ejaz, S. A.; Zargar, S.; Akhtar, N.; Aborode, A. T.; A Wani, T.; Batiha, G. E.-S.; Siddique, F.; Alqarni, M.; Akintola, A. A. Deep learning and structure-based virtual screening for drug discovery against NEK7: A novel target for the treatment of cancer. *Molecules* **2022**, *27*, 4098.
(14) Schaduangrat, N.; Lampa, S.; Simeon, S.; Gleeson, M. P.; Spjuth, O.; Nantasenamat, C. Towards reproducible computational drug discovery. *J. Cheminf.* **2020**, *12*, 9.
(15) Sliwoski, G.; Kothiwale, S.; Meiler, J.; Lowe, E. W., Jr Computational methods in drug discovery. *Pharmacol. Rev.* **2014**, *66*, 334−395.
(16) Choudhuri, S.; Yendluri, M.; Poddar, S.; Li, A.; Mallick, K.; Mallik, S.; Ghosh, B. Recent advancements in computational drug design algorithms through machine learning and optimization. *Kinases and Phosphatases* **2023**, *1*, 117−140.
(17) Rimac, H.; Grishina, M.; Potemkin, V. Use of the complementarity principle in docking procedures: A new approach for evaluating the correctness of binding poses. *J. Chem. Inf. Model.* **2021**, *61*, 1801−1813.
(18) Wang, Y.; Shaikh, S. A.; Tajkhorshid, E. Exploring transmembrane diffusion pathways with molecular dynamics. *Physiology* **2010**, *25*, 142−154.
(19) Van Westen, G. J.; Wegner, J. K.; Ijzerman, A. P.; Van Vlijmen, H. W.; Bender, A. Proteochemometric modeling as a tool to design selective compounds and for extrapolating to novel targets. *MedChemComm* **2011**, *2*, 16−30.
(20) Mak, K.-K.; Pichika, M. R. Artificial intelligence in drug development: present status and future prospects. *Drug discovery today* **2019**, *24*, 773−780.
(21) Zhu, H. Big data and artificial intelligence modeling for drug discovery. *Annu. Rev. Pharmacol. Toxicol.* **2020**, *60*, 573−589.
(22) Gupta, R.; Srivastava, D.; Sahu, M.; Tiwari, S.; Ambasta, R. K.; Kumar, P. Artificial intelligence to deep learning: machine intelligence approach for drug discovery. *Mol. Diversity* **2021**, *25*, 1315−1360.
(23) Patel, L.; Shukla, T.; Huang, X.; Ussery, D. W.; Wang, S. Machine learning methods in drug discovery. *Molecules* **2020**, *25*, 5277.
(24) Chen, H.; Engkvist, O.; Wang, Y.; Olivecrona, M.; Blaschke, T. The rise of deep learning in drug discovery. *Drug discovery today* **2018**, *23*, 1241−1250.
(25) Vattikuti, M. C. *Improving Drug Discovery and Development Using AI: Opportunities and Challenges*; Research-gate journal, 2024; Vol. *10*.
(26) Schneider, G. Automating drug discovery. *Nat. Rev. Drug Discovery* **2018**, *17*, 97−113.
(27) Bi, X.; Wang, Y.; Wang, J.; Liu, C. Machine Learning for Multi-Target Drug Discovery: Challenges and Opportunities in Systems Pharmacology. *Pharmaceutics* **2025**, *17*, 1186.
(28) Kleandrova, V. V.; Cordeiro, M. N. D.; Speck-Planche, A. Perturbation-Theory Machine Learning for Multi-Target Drug Discovery in Modern Anticancer Research. *Current Issues in Molecular Biology* **2025**, *47*, 301.
(29) Diéguez-Santana, K.; Casañola-Martin, G. M.; Green, J. R.; Rasulev, B.; González-Díaz, H. Predicting metabolic reaction networks with perturbation-theory machine learning (PTML) models. *Curr. Top. Med. Chem.* **2021**, *21*, 819−827.
(30) Kleandrova, V. V.; Cordeiro, M. N. D.; Speck-Planche, A. Perturbation-theory machine learning for mood disorders: virtual design of dual inhibitors of NET and SERT proteins. *BMC chemistry* **2025**, *19*, 2.
(31) Kleandrova, V. V.; Cordeiro, M. N. D.; Speck-Planche, A. Perturbation theory machine learning model for phenotypic early antineoplastic drug discovery: design of virtual anti-lung-cancer agents. *Appl. Sci.* **2024**, *14*, 9344.
(32) Kleandrova, V. V.; Speck-Planche, A. PTML modeling for pancreatic cancer research: in silico design of simultaneous multi-protein and multi-cell inhibitors. *Biomedicines* **2022**, *10*, 491.
(33) Speck-Planche, A.; Kleandrova, V. V.; Scotti, M. T. In silico drug repurposing for anti-inflammatory therapy: virtual search for dual inhibitors of caspase-1 and TNF-alpha. *Biomolecules* **2021**, *11*, 1832.
(34) Suay-García, B.; Climent, J.; Pérez-Gracia, M. T.; Falcó, A. A comprehensive update on the use of molecular topology applications for anti-infective drug discovery. *Expert Opinion on Drug Discovery* **2025**, *20*, 465−474.
(35) Segler, M. H.; Kogej, T.; Tyrchan, C.; Waller, M. P. Generating focused molecule libraries for drug discovery with recurrent neural networks. *ACS central science* **2018**, *4*, 120−131.

(36) Gupta, A.; Müller, A. T.; Huisman, B. J.; Fuchs, J. A.; Schneider, P.; Schneider, G. Generative recurrent networks for de novo drug design. *Molecular informatics* **2018**, *37*, 1700111.
(37) Grisoni, F.; Moret, M.; Lingwood, R.; Schneider, G. Bidirectional molecule generation with recurrent neural networks. *J. Chem. Inf. Model.* **2020**, *60*, 1175−1183.
(38) Olivecrona, M.; Blaschke, T.; Engkvist, O.; Chen, H. Molecular de-novo design through deep reinforcement learning. *J. Cheminf.* **2017**, *9*, 48.
(39) Chen, P.; Liu, W.; Hsieh, C.-Y.; Chen, G.; Zhang, S. Utilizing edge features in graph neural networks via variational information maximization. *arXiv* **2019**, arXiv:1906.05488 preprint.
(40) Honda, S.; Shi, S.; Ueda, H. R. Smiles transformer: Pre-trained molecular fingerprint for low data drug discovery. *arXiv* **2019**, arXiv:1911.04738 preprint.
(41) Adilov, S. Generative pre-training from molecules. *ChemRxiv* **2021**, 5fwjd This content is a preprint and has not been peer-reviewed.
(42) Gilmer, J.; Schoenholz, S. S.; Riley, P. F.; Vinyals, O.; Dahl, G. E. Neural message passing for quantum chemistry. In *International conference on machine learning*, 2017; pp 1263−1272.
(43) Hao, Z.; Lu, C.; Huang, Z.; Wang, H.; Hu, Z.; Liu, Q.; Chen, E.; Lee, C. A. S. G. N. An active semi-supervised graph neural network for molecular property prediction. In *Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining*, 2020; pp 731−752.
(44) Hu, W.; Liu, Y.; Chen, X.; Chai, W.; Chen, H.; Wang, H.; Wang, G. Deep learning methods for small molecule drug discovery: a survey. *IEEE Transactions on Artificial Intelligence* **2024**, *5*, 459−479.
(45) Blanco-Gonzalez, A.; Cabezon, A.; Seco-Gonzalez, A.; Conde-Torres, D.; Antelo-Riveiro, P.; Pineiro, A.; Garcia-Fandino, R. The role of AI in drug discovery: challenges, opportunities, and strategies. *Pharmaceuticals* **2023**, *16*, 891.
(46) Bran, A. M.; Schwaller, P. *Drug Development Supported by Informatics*; Springer, 2024; pp 143−163.
(47) R, N.; Khan, S. B.; kumar, A. V.; Alojail, M.; Alojail, M.; Sangwan, S. R.; Saraee, M. Enhancing drug discovery and patient care through advanced analytics with the power of NLP and machine learning in pharmaceutical data interpretation. *SLAS Technology* **2025**, *31*, 100238.
(48) Bhatnagar, R.; Sardar, S.; Beheshti, M.; Podichetty, J. T. How can natural language processing help model informed drug development?: a review. *JAMIA Open* **2022**, *5*, ooac043.
(49) McEntire, R.; Szalkowski, D.; Butler, J.; Kuo, M. S.; Chang, M.; Chang, M.; Freeman, D.; et al. Application of an automated natural language processing (NLP) workflow to enable federated search of external biomedical content in drug discovery and development. *Drug Discovery Today* **2016**, *21*, 826−835.
(50) Öztürk, H.; Özgür, A.; Schwaller, P.; Laino, T.; Ozkirimli, E. Exploring chemical space using natural language processing methodologies for drug discovery. *Drug Discovery Today* **2020**, *25*, 689−705.
(51) Vaswani, A. Attention is all you need. In *Advances in Neural Information Processing Systems*, 2017.
(52) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. *nature* **2021**, *596*, 583−589.
(53) Varadi, M.; Anyango, S.; Deshpande, M.; Nair, S.; Natassia, C.; Yordanova, G.; Yuan, D.; et al. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* **2022**, *50*, D439−D444.
(54) An, J.; Ding, W.; Lin, C. ChatGPT. tackle the growing carbon footprint of generative AI. *Nature* **2023**, *615*, 586.
(55) Winter, B.; Winter, C.; Schilling, J.; Bardow, A. A smile is all you need: predicting limiting activity coefficients from SMILES with natural language processing. *Digital Discovery* **2022**, *1*, 859−869.
(56) Ding, Y.; Tang, J.; Guo, F. Identification of drug-target interactions via multiple information integration. *Information Sciences* **2017**, *418*, 546−560.
(57) Li, A. P. *Drug-drug interactions in pharmaceutical development*; Wiley-Interscience, 2008.
(58) Karimi, S.; Wang, C.; Metke-Jimenez, A.; Gaire, R.; Paris, C. Text and data mining techniques in adverse drug reaction detection. *ACM Computing Surveys* **2015**, *47*, 1−39.
(59) Dou, M.; Tang, J.; Tiwari, P.; Ding, Y.; Guo, F. Drug−Drug Interaction Relation Extraction Based on Deep Learning: A Review. *ACM Computing Surveys* **2024**, *56*, 1−33.
(60) Qiu, Y.; Zhang, Y.; Deng, Y.; Liu, S.; Zhang, W. A comprehensive review of computational methods for drug-drug interaction detection. *IEEE/ACM transactions on computational biology and bioinformatics* **2022**, *19*, 1968−1985.
(61) Zhang, T.; Leng, J.; Liu, Y. Deep learning for drug−drug interaction extraction from the literature: a review. *Briefings in bioinformatics* **2020**, *21*, 1609−1627.
(62) Han, K.; Cao, P.; Wang, Y.; Xie, F.; Ma, J.; Yu, M.; Wang, J.; Xu, Y.; Zhang, Y.; Wan, J. A review of approaches for predicting drug−drug interactions based on machine learning. *Frontiers in pharmacology* **2022**, *12*, 814858.
(63) Bagherian, M.; Sabeti, E.; Wang, K.; Sartor, M. A.; Nikolovska-Coleska, Z.; Najarian, K. Machine learning approaches and databases for prediction of drug−target interaction: a survey paper. *Briefings in bioinformatics* **2021**, *22*, 247−269.
(64) Kim, J.; Park, S.; Min, D.; Kim, W. Comprehensive survey of recent drug discovery using deep learning. *Int. J. Mol. Sci.* **2021**, *22*, 9983.
(65) Radwan, Y.; Gouda, k.; Abdelbaky, I.; arafa, m. A comprehensive survey explores Drug-Drug interaction prediction using Machine-Learning techniques. *Benha Journal of Applied Sciences* **2024**, *9*, 13−21.
(66) Budgen, D.; Brereton, P. Performing systematic literature reviews in software engineering. In *Proceedings of the 28th international conference on Software engineering*, 2006; pp 1051−1052.
(67) Miller, T.; Geva, A.; Dligach, D. Extracting Adverse Drug Event Information with Minimal Engineering. In *Proceedings of the 2nd Clinical Natural Language Processing Workshop*, 2019; pp 22−27. Minneapolis, Minnesota, USA.
(68) Herman Bernardim Andrade, G.; Nishiyama, T.; Fujimaki, T.; Yada, S.; Wakamiya, S.; Takagi, M.; Kato, M.; Miyashiro, I.; Aramaki, E. *Assessing Adverse Drug Event Extraction Performance on Real-World Breast Cancer Data Through Domain Adaptation*. Tomohiro and Fujimaki, Takako and Yada, Shuntaro and Wakamiya, Shoko and Takagi, Mari and Kato, Mizuki and Miyashiro, Isao and Aramaki, Eiji, Assessing Adverse Drug Event Extraction Performance on Real-World Breast Cancer Data Through Domain Adaptation 2023.
(69) Syrowatka, A.; Song, W.; Amato, M. G.; Foer, D.; Edrees, H.; Co, Z.; Kuznetsova, M.; Dulgarian, S.; Seger, D. L.; Simona, A.; et al. Key use cases for artificial intelligence to reduce the frequency of adverse drug events: a scoping review. *Lancet Digital Health* **2022**, *4*, e137−e148.
(70) Salman, M.; Munawar, H. S.; Latif, K.; Akram, M. W.; Khan, S. I.; Ullah, F. Big data management in drug−drug interaction: a modern deep learning approach for smart healthcare. *Big Data and Cognitive Computing* **2022**, *6*, 30.
(71) Lenz, W. A short history of thalidomide embryopathy. *Teratology* **1988**, *38*, 203−215.
(72) Shafique, S. Thalidomidean overview and the species-specific teratogenicity. *Res. J. Congenit Dis* **2019**, *2*, 1.
(73) Bhattacharya, S. *Safe and effective medicines for use in pregnancy: a call to action. Birmingham: Birmingham Health Partners Centre for Regulatory Science and Innovation*; University of Birmingham, 2021.
(74) Wei, J.; Hu, T.; Dai, J.; Wang, Z.; Han, P.; Huang, W. Research on named entity recognition of adverse drug reactions based on NLP and deep learning. *Front. Pharmacol* **2023**, *14*, 1121796.
(75) Noah, B. A. Adverse Drug Reactions: Harnessing Experiential Data to Promote Patient Welfare. *Cath. UL Rev.* **2000**, *49*, 449.
(76) Steinzor, R.; Clune, M. *The Hidden Lesson of the Vioxx Fiasco: Reviving a Hollow FDA*; Center for Progressive Reform, 2005; .
(77) Jha, N.; Rathore, D. S.; Shankar, P. R.; Gyawali, S. Pharmacovigilance knowledge among patients at a teaching hospital in Lalitpur district, Nepal. *Journal of clinical and diagnostic research: JCDR* **2014**, *8*, 32.

https://doi.org/10.1021/acsomega.5c04997
ACS Omega 2025, 10, 61048−61075

(78) Soni, R.; Kesari, B. A review on pharmacovigilance. *Int. J. Pharm. Sci. Rev. Res.* **2014**, *26*, 237−241.
(79) Schatz, S.; Weber, R. J. Adverse drug reactions. *Pharmacy Practice* **2015**, *1*, 16−23.
(80) McMaster, C.; Chan, J.; Liew, D. F.; Su, E.; Frauman, A. G.; Chapman, W. W.; Pires, D. E. Developing a deep learning natural language processing algorithm for automated reporting of adverse drug reactions. *J. Biomed. Inf.* **2023**, *137*, 104265.
(81) Rewale, K. S.; Tuwar, A. B.; Salve, M. T. A Brief Review on Pharmacovigilance. *International Journal of Humanities Social Science and Management (IJHSSM)* **2023**, *3*, 156.
(82) White, T. J.; Arakelian, A.; Rho, J. P. Counting the costs of drug-related adverse events. *Pharmacoeconomics* **1999**, *15*, 445−458.
(83) Sultana, J.; Cutroneo, P.; Trifirò, G. Clinical and economic burden of adverse drug reactions. *Journal of Pharmacology and Pharmacotherapeutics* **2013**, *4*, S73−S77.
(84) Ehsani, J. P.; Jackson, T.; Duckett, S. J. The incidence and cost of adverse events in Victorian hospitals 2003−04. *Medical Journal of Australia* **2006**, *184*, 551−555.
(85) Moura, C. S. d.; Acurcio, F. d. A.; Belo, N. d. O. Drug-drug interactions associated with length of stay and cost of hospitalization. *J. Pharm. Pharm. Sci.* **2009**, *12*, 266−272.
(86) Percha, B.; Altman, R. B. Informatics confronts drug−drug interactions. *Trends in pharmacological sciences* **2013**, *34*, 178−184.
(87) Becker, M. L.; Kallewaard, M.; Caspers, P. W.; Visser, L. E.; Leufkens, H. G.; Stricker, B. H. Hospitalisations and emergency department visits due to drug−drug interactions: a literature review. *Pharmacoepidemiology and drug safety* **2007**, *16*, 641−651.
(88) Bernstein, A.; Bilheimer, L. T.; Makuc, D. M. *Health, United States, 2011: With special feature on socioeconomic status and health*. DHHS publication; no. 2012−1232 2012,.
(89) Gallelli, L.; Ferreri, G.; Colosimo, M.; Pirritano, D.; Guadagnino, L.; Pelaia, G.; Maselli, R.; De Sarro, G. Adverse drug reactions to antibiotics observed in two pulmonology divisions of Catanzaro, Italy: a six-year retrospective study. *Pharmacological research* **2002**, *46*, 395−400.
(90) Gallelli, L.; Ferreri, G.; Colosimo, M.; Pirritano, D.; Flocco, M.; Pelaia, G.; Maselli, R.; De Sarro, G. Retrospective analysis of adverse drug reactions to bronchodilators observed in two pulmonary divisions of Catanzaro, Italy. *Pharmacological research* **2003**, *47*, 493−499.
(91) Gallelli, L.; Colosimo, M.; Tolotta, G. A.; Falcone, D.; Luberto, L.; Curto, L. S.; Rende, P.; Mazzei, F.; Marigliano, N. M.; De Sarro, G.; et al. Prospective randomized double-blind trial of racecadotril compared with loperamide in elderly people with gastroenteritis living in nursing homes. *European journal of clinical pharmacology* **2010**, *66*, 137−144.
(92) Palleria, C.; Di Paolo, A.; Giofrè, C.; Caglioti, C.; Leuzzi, G.; Siniscalchi, A.; De Sarro, G.; Gallelli, L. Pharmacokinetic drug-drug interaction and their implication in clinical management. *Journal of research in medical sciences: the official journal of Isfahan University of Medical Sciences* **2013**, *18*, 601.
(93) Siniscalchi, A.; Gallelli, L.; Avenoso, T.; Squillace, A.; De Sarro, G. Effects of carbamazepine/oxycodone coadministration in the treatment of trigeminal neuralgia. *Annals of Pharmacotherapy* **2011**, *45*, No. e33.
(94) Canu, B.; Fioravanti, A.; Orlandi, P.; Di Desidero, T.; Alì, G.; Fontanini, G.; Di Paolo, A.; Del Tacca, M.; Danesi, R.; Bocci, G. Irinotecan synergistically enhances the antiproliferative and proapoptotic effects of axitinib in vitro and improves its anticancer activity in vivo. *Neoplasia* **2011**, *13*, 217.
(95) KafiKang, M.; Hendawi, A. Drug-drug interaction extraction from biomedical text using relation BioBERT with BLSTM. *Machine Learning and Knowledge Extraction* **2023**, *5*, 669−683.
(96) Machado, J.; Rodrigues, C.; Sousa, R.; Gomes, L. M. Drug−drug interaction extraction-based system: An natural language processing approach. *Expert Systems* **2023**, *42*, No. e13303.
(97) Bui, Q. Relation extraction methods for biomedical literature. Ph.D. Thesis, Universiteit van Amsterdam, 2012.
(98) Sun, W.; Cai, Z.; Li, Y.; Liu, F.; Fang, S.; Wang, G. Data processing and text mining technologies on electronic medical records: a review. *Journal of healthcare engineering* **2018**, *2018*, 1−9.
(99) Talib, R.; Kashif, M.; Ayesha, S.; Fatima, F. Text mining: techniques, applications and issues. *International journal of advanced computer science and applications* **2016**, *7*, 414−418.
(100) Luque, C.; Luna, J. M.; Luque, M.; Ventura, S. An advanced review on text mining in medicine. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* **2019**, *9*, No. e1302.
(101) Zirkle, J.; Han, X.; Racz, R.; Samieegohar, M.; Chaturbedi, A.; Mann, J.; Chakravartula, S.; Li, Z. Deep learning-enabled natural language processing to identify directional pharmacokinetic drug−drug interactions. *BMC bioinformatics* **2023**, *24*, 413.
(102) Segura-Bedmar, I.; Martínez, P.; Herrero-Zazo, M. Semeval-2013 task 9: Extraction of drug-drug interactions from biomedical texts (ddiextraction 2013). In *Second Joint Conference on Lexical and Computational Semantics (* SEM), Volume 2: Proceedings of the Seventh International Workshop on Semantic Evaluation (SemEval 2013)*, 2013; pp 341−350.
(103) Thomas, P.; Neves, M.; Rocktäschel, T.; Leser, U. WBI-DDI: drug-drug interaction extraction using majority voting. In *Second Joint Conference on Lexical and Computational Semantics (* SEM), Volume 2: Proceedings of the Seventh International Workshop on Semantic Evaluation (SemEval 2013)*, 2013; pp 628−635.
(104) Young, T.; Hazarika, D.; Poria, S.; Cambria, E. Recent trends in deep learning based natural language processing. *ieee Computational intelligenCe magazine* **2018**, *13*, 55−75.
(105) Choudhary, K.; DeCost, B.; Chen, C.; Jain, A.; Tavazza, F.; Cohn, R.; Park, C. W.; Choudhary, A.; Agrawal, A.; Billinge, S. J.; et al. Recent advances and applications of deep learning methods in materials science. *npj Comput. Mater.* **2022**, *8*, 59.
(106) Ciaburro, G.; Venkateswaran, B. *Neural Networks with R: Smart models using CNN, RNN, deep learning, and artificial intelligence principles*; Packt Publishing Ltd, 2017.
(107) Otter, D. W.; Medina, J. R.; Kalita, J. K. A survey of the usages of deep learning for natural language processing. *IEEE transactions on neural networks and learning systems* **2021**, *32*, 604−624.
(108) Park, K. A review of computational drug repurposing. *Translational and clinical pharmacology* **2019**, *27*, 59.
(109) Bumb, J. M.; Enning, F.; Leweke, F. M. Drug repurposing and emerging adjunctive treatments for schizophrenia. *Expert Opinion on Pharmacotherapy* **2015**, *16*, 1049−1067.
(110) Pushpakom, S.; Iorio, F.; Eyers, P. A.; Escott, K. J.; Hopper, S.; Wells, A.; Doig, A.; Guilliams, T.; Latimer, J.; McNamee, C.; et al. Drug repurposing: progress, challenges and recommendations. *Nat. Rev. Drug Discovery* **2019**, *18*, 41−58.
(111) Eslami Manoochehri, H.; Nourani, M. Drug-target interaction prediction using semi-bipartite graph model and deep learning. *BMC bioinformatics* **2020**, *21*, 248.
(112) Talevi, A.; Bellera, C. L. Challenges and opportunities with drug repurposing: finding strategies to find alternative uses of therapeutics. *Expert Opinion on Drug Discovery* **2020**, *15*, 397−401.
(113) Croset, S. *Drug repositioning and indication discovery using description logics*, University of Cambridge, 2014. PhD Thesis.
(114) Tanoli, Z.; Vähä-Koskela, M.; Aittokallio, T. Artificial intelligence, machine learning, and drug repurposing in cancer. *Expert Opinion on Drug Discovery* **2021**, *16*, 977−989.
(115) Renaud, R.; Xuereb, H. Erectile-dysfunction therapies. *Nat. Rev. Drug Discovery* **2002**, *1*, 663.
(116) Wu, Z.; Li, W.; Liu, G.; Tang, Y. Network-based methods for prediction of drug-target interactions. *Frontiers in pharmacology* **2018**, *9*, 1134.
(117) Li, F.; Zhang, Z.; Guan, J.; Zhou, S. Effective drug−target interaction prediction with mutual interaction neural network. *Bioinformatics* **2022**, *38*, 3582−3589.
(118) Yue, Y.; He, S. DTI-HeNE: a novel method for drug-target interaction prediction based on heterogeneous network embedding. *BMC bioinformatics* **2021**, *22*, 418.

https://doi.org/10.1021/acsomega.5c04997
ACS Omega 2025, 10, 61048−61075

(119) Peng, Y.; Wang, J.; Wu, Z.; Zheng, L.; Wang, B.; Liu, G.; Li, W.; Tang, Y. MPSM-DTI: prediction of drug−target interaction via machine learning based on the chemical structure and protein sequence. *Digital Discovery* **2022**, *1*, 115−126.
(120) Kapetanovic, I. Computer-aided drug discovery and development (CADDD): in silico-chemico-biological approach. *Chemico-biological interactions* **2008**, *171*, 165−176.
(121) Lee, I.; Keum, J.; Nam, H. DeepConv-DTI: Prediction of drug-target interactions via deep learning with convolution on protein sequences. *PLoS computational biology* **2019**, *15*, No. e1007129.
(122) Zhao, T.; Hu, Y.; Valsdottir, L. R.; Zang, T.; Peng, J. Identifying drug−target interactions based on graph convolutional network and deep neural network. *Briefings in bioinformatics* **2021**, *22*, 2141−2150.
(123) Liu, X.; Jiang, H.; Li, H. SHAFTS: a hybrid approach for 3D molecular similarity calculation. 1. Method and assessment of virtual screening. *J. Chem. Inf. Model.* **2011**, *51*, 2372−2385.
(124) Gfeller, D.; Michielin, O.; Zoete, V. Shaping the interaction landscape of bioactive molecules. *Bioinformatics* **2013**, *29*, 3073−3079.
(125) Gong, J.; Cai, C.; Liu, X.; Ku, X.; Jiang, H.; Gao, D.; Li, H. ChemMapper: a versatile web server for exploring pharmacology and chemical structure association based on molecular 3D similarity method. *Bioinformatics* **2013**, *29*, 1827−1829.
(126) Daina, A.; Michielin, O.; Zoete, V. SwissTargetPrediction: updated data and new features for efficient prediction of protein targets of small molecules. *Nucleic acids research* **2019**, *47*, W357−W364.
(127) Li, J.; Zheng, S.; Chen, B.; et al. A survey of current trends in computational drug repositioning. *Briefings in Bioinformatics* **2016**, *17*, 2−12.
(128) Nascimento, A.; Prudêncio, R. B. C.; Costa, I. A multiple kernel learning algorithm for drug-target interaction prediction. *BMC Bioinf.* **2016**, *17*, 46.
(129) Xia, Z.; Wu, L.-Y.; Zhou, X.; Wong, S. T. Semi-supervised drug-protein interaction prediction from heterogeneous biological spaces. *BMC Systems Biology* **2010**, *4*, S6.
(130) Bleakley, K.; Yamanishi, Y. Supervised prediction of drugtarget interactions using bipartite local models. *Bioinformatics* **2009**, *25*, 2397−2403.
(131) Boutet, E.; Lieberherr, D.; Tognolli, M.; Schneider, M.; Bairoch, A. *Plant Bioinformatics: Methods and Protocols*; Humana Press: Totowa, NJ, 2007; pp 89−112.
(132) Boutet, E.; Lieberherr, D.; Tognolli, M.; Schneider, M.; Bansal, P.; Bridge, A. J.; Poux, S.; Bougueleret, L.; Xenarios, I. *Plant Bioinformatics: Methods and Protocols*; Humana Press: New York, NY, 2016; pp 23−54.
(133) Martin, A. C. Mapping PDB Chains to UniProtKB Entries. *Bioinformatics* **2005**, *21*, 4297−4301.
(134) Schneider, M.; Bairoch, A.; Wu, C. H.; Apweiler, R. Plant Protein Annotation in the UniProt Knowledgebase. *Plant Physiol.* **2005**, *138*, 59−66.
(135) Breuza, L.; Poux, S.; Estreicher, A.; Famiglietti, M. L.; Magrane, M.; Tognolli, M.; Bridge, A.; Baratin, D.; Redaschi, N.; The UniProt Consortium. The UniProtKB Guide to the Human Proteome. *Database* **2016**, *2016*, bav120.
(136) MacDougall, A.; Volynkin, V.; Saidi, R.; Poggioli, D.; Zellner, H.; Hatton-Ellis, E.; Joshi, V. e. a.; O'Donovan, C.; Orchard, S.; Auchincloss, A. H.; et al. UniRule: A Unified Rule Resource for Automatic Annotation in the UniProt Knowledgebase. *Bioinformatics* **2020**, *36*, 4643−4648.
(137) Asija, R.; Nallusamy, R. Healthcare SaaS Based on a Data Model with Built-in Security and Privacy. *International Journal of Cloud Applications and Computing (IJCAC)* **2016**, *6*, 1−14.
(138) The UniProt Consortium. UniProt: a hub for protein information. *Nucleic acids research* **2015**, *43*, D204−D212.
(139) Magrane, M.; Consortium, U. UniProt Knowledgebase: a hub of integrated protein data. *Database* **2011**, *2011*, bar009.
(140) Denny, P.; The UniProt Consortium; et al. UniProt: the Universal protein knowledgebase in 2025. *Nucleic Acids Res.* **2025**, *53*, D609−D617.
(141) Masic, I.; Ferhatovica, A. Review of Most Important Biomedical Databases for Searching of Biomedical Scientific Literature. *Donald School Journal of Ultrasound in Obstetrics and Gynecology* **2012**, *6*, 343−361.
(142) Martenot, V.; Masdeu, V.; Cupe, J.; Gehin, F.; Blanchon, M.; Dauriat, J.; Horst, A.; Renaudin, M.; Girard, P.; Zucker, J.-D. LiSA: an assisted literature search pipeline for detecting serious adverse drug events with deep learning. *BMC medical informatics and decision making* **2022**, *22*, 338.
(143) Wishart, D. S.; Feunang, Y. D.; Guo, A. C.; Lo, E. J.; Marcu, A.; Grant, J. R.; Sajed, T.; Johnson, D.; Li, C.; Sayeeda, Z.; et al. DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic acids research* **2018**, *46*, D1074−D1082.
(144) Knox, C.; Wilson, M.; Klinger, C. M.; Franklin, M.; Oler, E.; Wilson, A.; Pon, A.; Cox, J.; Chin, N. E.; Strawbridge, S. A.; et al. DrugBank 6.0: the DrugBank knowledgebase for 2024. *Nucleic acids research* **2024**, *52*, D1265−D1275.
(145) Liu, T.; Hwang, L.; Burley, S. K.; Nitsche, C. I.; Southan, C.; Walters, W. P.; Gilson, M. K. BindingDB in 2024: a FAIR knowledgebase of protein-small molecule binding data. *Nucleic acids research* **2025**, *53*, D1633−D1644.
(146) Gilson, M. K.; Liu, T.; Baitaluk, M.; Nicola, G.; Hwang, L.; Chong, J. BindingDB in 2015: a public database for medicinal chemistry, computational chemistry and systems pharmacology. *Nucleic acids research* **2016**, *44*, D1045−D1053.
(147) Chen, X.; Liu, M.; Gilson, M. K. BindingDB: a web-accessible molecular recognition database. *Combinatorial chemistry & high throughput screening* **2001**, *4*, 719−725.
(148) Tian, Y.; Yi, J.; Wang, N.; Wu, C.; Peng, J.; Liu, S.; Yang, G.; Cao, D. DDInter 2.0: an enhanced drug interaction resource with expanded data coverage, new interaction types, and improved user interface. *Nucleic Acids Res.* **2025**, *53*, D1356−D1362.
(149) Fu, L.; Shi, S.; Yi, J.; Wang, N.; He, Y.; Wu, Z.; Peng, J.; Deng, Y.; Wang, W.; Wu, C.; et al. ADMETlab 3.0: an updated comprehensive online ADMET prediction platform enhanced with broader coverage, improved performance, API functionality and decision support. *Nucleic acids research* **2024**, *52*, W422−W431.
(150) Nebert, D. W.; Russell, D. W. Clinical importance of the cytochromes P450. *Lancet* **2002**, *360*, 1155−1162.
(151) Wu, H.; Xing, Y.; Ge, W.; Liu, X.; Zou, J.; Zhou, C.; Liao, J. Drug-drug interaction extraction via hybrid neural networks on biomedical literature. *J. Biomed. Inf.* **2020**, *106*, 103432.
(152) U.S. National Institute of Standards and Technolog; *Technology Text Analysis Conference (TAC)* 2018. https://tac.nist.gov/2018/index.html, 2018.
(153) Goodwin, T. R.; Demner-Fushman, D.; Fung, K. W.; Do, P. Overview of the TAC 2019 Track on Drug-Drug Interaction Extraction from Drug Labels. In *Text Analysis Conference (TAC)*, 2019.
(154) Tran, T.; Kavuluru, R.; Kilicoglu, H. A Multi-Task Learning Framework for Extracting Drugs and Their Interactions from Drug Labels. *arXiv* **2019**, arXiv:1905.07464.
(155) Dong, F.; Guo, W.; Liu, J.; Patterson, T. A.; Hong, H. BERT-based language model for accurate drug adverse event extraction from social media: implementation, evaluation, and contributions to pharmacovigilance practices. *Frontiers in Public Health* **2024**, *12*, 1392180.
(156) Tayefi, M.; Ngo, P.; Chomutare, T.; Dalianis, H.; Salvi, E.; Budrionis, A.; Godtliebsen, F. Challenges and Opportunities Beyond Structured Data in Analysis of Electronic Health Records. *Wiley Interdisciplinary Reviews: Computational Statistics* **2021**, *13*, No. e1549.
(157) Mahajan, D.; Liang, J. J.; Tsou, C.-H. Toward understanding clinical context of medication change events in clinical narratives. In *AMIA Annual Symposium Proceedings.*, 2021; p 833.
(158) Salek Faramarzi, N.; Patel, M.; Bandarupally, S. H.; Banerjee, R. Context-aware Medication Event Extraction from Unstructured Text. In *Proceedings of the 5th Clinical Natural Language Processing Workshop*, **2023**; pp 86−95.
(159) Chen, T.-Y.; Aquino, J. A.; Chiu, Y.-W.; Yeh, W.-C.; Chang, Y.-C. Understanding the Clinical Context of Medication Change Events in

https://doi.org/10.1021/acsomega.5c04997
ACS Omega 2025, 10, 61048−61075

Clinical Narratives using Pre-trained Clinical Language Models. In *Proceedings of the 2023 7th International Conference on Medical and Health Informatics*, 2023; pp 98−103.
(160) Ramachandran, G. K.; Lybarger, K.; Liu, Y.; Mahajan, D.; Liang, J. J.; Tsou, C.-H.; Yetisgen, M.; Uzuner, A.~. Extracting Medication Changes in Clinical Narratives Using Pre-trained Language Models. *Journal of Biomedical Informatics* **2023**, *139*, 104302.
(161) Uzuner, A.~; Stubbs, A. Practical Applications for Natural Language Processing in Clinical Research: The 2014 i2b2/UTHealth Shared Tasks. *Journal of Biomedical Informatics* **2015**, *58*, S1.
(162) Davis, M. I.; Hunt, J. P.; Herrgard, S.; Ciceri, P.; Wodicka, L. M.; Pallares, G.; Hocker, M.; Treiber, D. K.; Zarrinkar, P. P. Comprehensive analysis of kinase inhibitor selectivity. *Nature biotechnology* **2011**, *29*, 1046−1051.
(163) Liu, B. Drug-target affinity prediction method based on consistent expression of heterogeneous data. In *2022 International Conference on Informatics, Networking and Computing (ICINC)*, 2022; pp 116−121.
(164) De Souza, J. G.; Fernandes, M. A.; de Melo Barbosa, R. A novel deep neural network technique for drug−target interaction. *Pharmaceutics* **2022**, *14*, 625.
(165) Tang, J.; Szwajda, A.; Shakyawar, S.; Xu, T.; Hintsanen, P.; Wennerberg, K.; Aittokallio, T. Making sense of large-scale kinase inhibitor bioactivity data sets: a comparative and integrative analysis. *J. Chem. Inf. Model.* **2014**, *54*, 735−743.
(166) Edgar, R.; Barrett, T. NCBI GEO Standards and Services for Microarray Data. *Nat. Biotechnol.* **2006**, *24*, 1471−1472.
(167) Barrett, T.; Wilhite, S. E.; Ledoux, P.; Evangelista, C.; Kim, I. F.; Tomashevsky, M.; Marshall, K. A.; Phillippy, K. H.; Sherman, P. M.; Holko, M.; et al. NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Res.* **2012**, *41*, D991−D995.
(168) Williams, G.; Azim, K. A Predictive Model for Compound-Protein Interactions Based on Concatenated Vectorization. *bioRxiv* **2024**, 616275.
(169) International Human Genome Sequencing Consortium. Finishing the Euchromatic Sequence of the Human Genome. *Nature* **2004**, *431*, 931−945.
(170) The Chimpanzee Sequencing and Analysis Consortium. Initial Sequence of the Chimpanzee Genome and Comparison with the Human Genome. *Nature* **2005**, *437*, 69−87.
(171) The 1000 Genomes Project Consortium. A Map of Human Genome Variation from Population Scale Sequencing. *Nature* **2010**, *467*, 1061−1073.
(172) Lander, E. S.; Linton, L. M.; Birren, B.; Nusbaum, C.; Zody, M. C.; Baldwin, J.; Devon, K.; Dewar, K.; Doyle, M.; FitzHugh, W.; et al. Initial sequencing and analysis of the human genome. *Nature* **2001**, *409*, 860−921.
(173) Sutphin, C.; Lee, K.; Yepes, A. J.; Uzuner, O.·.; McInnes, B. T. Adverse drug event detection using reason assignments in FDA drug labels. *J. Biomed. Inf.* **2020**, *110*, 103552.
(174) Gurulingappa, H.; Rajput, A. M.; Roberts, A.; Fluck, J.; Hofmann-Apitius, M.; Toldo, L. Development of a benchmark corpus to support the automatic extraction of drug-related adverse effects from medical case reports. *J. Biomed. Inf.* **2012**, *45*, 885−892.
(175) Wang, K.; Fu, X.; Liu, Y.; Chen, W.; Chen, J. PTDA: Improving Drug-Drug Interaction Extraction from Biomedical Literature Based on Prompt Tuning and Data Augmentation. *IAENG International Journal of Computer Science* **2024**, 518.
(176) Shi, Y.; He, M.; Chen, J.; Han, F.; Cai, Y. SubGE-DDI: A new prediction model for drug-drug interaction established through biomedical texts and drug-pairs knowledge subgraph enhancement. *PLOS Computational Biology* **2024**, *20*, No. e1011989.
(177) Guo, X.; Song, D.; Yang, F. Enhanced Attention-Driven Dynamic Graph Convolutional Network for Extracting Drug-Drug Interaction. *Big Data Mining and Analytics* **2025**, *8*, 257−271.
(178) Zaikis, D.; Vlahavas, I. TransformDDI: The Transformer-Based Joint Multi-Task Model for End-to-End Drug-Drug Interaction Extraction. *IEEE Journal of Biomedical and Health Informatics* **2025**, *29*, 3045−3056.
(179) Hassan, N. A.; Seoud, R. A. A.; Salem, D. A. Bridging the Gap: A Hybrid Approach to Medical Relation Extraction Using Pretrained Language Models and Traditional Machine Learning. *Journal of Advances in Information Technology* **2024**, *15*, 723−734.
(180) De Vito, G.; Ferrucci, F.; Angelakis, A. LLMs for Drug-Drug Interaction Prediction: A Comprehensive Comparison. *arXiv* **2025**, arXiv:2502.06890.
(181) Deng, Y.; Xu, X.; Qiu, Y.; Xia, J.; Zhang, W.; Liu, S. A multimodal deep learning framework for predicting drug−drug interaction events. *Bioinformatics* **2020**, *36*, 4316−4322.
(182) Yang, Z.; Tong, K.; Jin, S.; Wang, S.; Yang, C.; Jiang, F. CNN-Siam: multimodal siamese CNN-based deep learning approach for drug−drug interaction prediction. *BMC bioinformatics* **2023**, *24*, 110.
(183) Zhang, Y.; Cui, S.; Gao, H. Adverse drug reaction detection on social media with deep linguistic features. *J. Biomed. Inf.* **2020**, *106*, 103437.
(184) Herrero-Zazo, M.; Segura-Bedmar, I.; Martínez, P.; Declerck, T. The DDI corpus: An annotated corpus with pharmacological substances and drug−drug interactions. *J. Biomed. Inf.* **2013**, *46*, 914−920.
(185) Krallinger, M.; Leitner, F.; Rabal, O.; Vazquez, M.; Oyarzabal, J.; Valencia, A. CHEMDNER: The drugs and chemical names extraction challenge. *J. Cheminf.* **2015**, *7*, S1.
(186) Niu, D.; Zhang, L.; Zhang, B.; Zhang, Q.; Li, Z. DAS-DDI: A dual-view framework with drug association and drug structure for drug-drug interaction prediction. *Journal of Biomedical Informatics* **2024**, *156*, 104672.
(187) Hu, W.; Fey, M.; Zitnik, M.; Dong, Y.; Ren, H.; Liu, B.; Catasta, M.; Leskovec, J. Open graph benchmark: Datasets for machine learning on graphs. In *Advances in neural information processing systems*, 2020; pp 22118−22133.
(188) Su, X.; Zhao, B.; Li, G.; Zhang, J.; Hu, P.; You, Z.; Hu, L. Knowledge graph neural network with spatial-aware capsule for drug-drug interaction prediction. *IEEE journal of biomedical and health informatics* **2025**, *29*, 1771−1781.
(189) Ogata, H.; Goto, S.; Sato, K.; Fujibuchi, W.; Bono, H.; Kanehisa, M. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic acids research* **1999**, *27*, 29−34.
(190) Luo, Y.; Zhao, X.; Zhou, J.; Yang, J.; Zhang, Y.; Kuang, W.; Peng, J.; Chen, L.; Zeng, J. A network integration approach for drug-target interaction prediction and computational drug repositioning from heterogeneous information. *Nat. Commun.* **2017**, *8*, 573.
(191) Zhang, C.; Sun, J.; Xing, L.; Zhang, L.; Cai, H.; Guo, M. MHANDTI: Drug-Target Interaction Prediction Model Based on Heterogeneous Graph Multi-Hop Attention Networks. In *2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2024; pp 475−478.
(192) Zhu, Y.; Ning, C.; Zhang, N.; Wang, M.; Zhang, Y. GSRF-DTI a framework for drug-target interaction prediction based on a drug-target pair network and representation learning on a large graph. *BMC biology* **2024**, *22*, 156.
(193) Song, W.; Xu, L.; Han, C.; Tian, Z.; Zou, Q. Drug−target interaction predictions with multi-view similarity network fusion strategy and deep interactive attention mechanism. *Bioinformatics* **2024**, *40*, btae346.
(194) Wang, Y.; Yin, Z. Drug−target interaction prediction through fine-grained selection and bidirectional random walk methodology. *Sci. Rep.* **2024**, *14*, 18104.
(195) Ahmed, K. T.; Ansari, M. I.; Zhang, W. DTI-LM language model powered drug−target interaction prediction. *Bioinformatics* **2024**, *40*, btae533.
(196) Dehghan, A.; Abbasi, K.; Razzaghi, P.; Banadkuki, H.; Gharaghani, S. CCL-DTI: contributing the contrastive loss in drug−target interaction prediction. *BMC bioinformatics* **2024**, *25*, 48.
(197) Zhang, R.; Wang, Z.; Wang, X.; Meng, Z.; Cui, W. MHTAN-DTI Metapath-based hierarchical transformer and attention network for drug−target interaction prediction. *Briefings in Bioinformatics* **2023**, *24*, bbad079.

https://doi.org/10.1021/acsomega.5c04997
ACS Omega 2025, 10, 61048−61075

(198) Zitnik, M.; Agrawal, M.; Leskovec, J. Modeling polypharmacy side effects with graph convolutional networks. *Bioinformatics* **2018**, *34*, i457−i466.
(199) Liu, S.; Liu, Y.; Xu, H.; Xia, J.; Li, S. Z. SP-DTI Subpocket-Informed Transformer for Drug-Target Interaction Prediction. *Bioinformatics* **2025**, *41*, btaf011.
(200) Lee, J.; Jun, D. W.; Song, I.; Kim, Y. DLM-DTI: a dual language model for the prediction of drug-target interaction with hint-based learning. *Journal of Cheminf.* **2024**, *16*, 14.
(201) Peng, L.; Liu, X.; Yang, L.; Liu, L.; Bai, Z.; Chen, M.; Lu, X.; Nie, L. BINDTI: a bi-directional intention network for drug-target interaction identification based on attention mechanisms. *IEEE Journal of Biomedical and Health Informatics* **2025**, *29*, 1602−1612.
(202) Talo, M.; Bozdag, S. Top-DTI: Integrating Topological Deep Learning and Large Language Models for Drug Target Interaction Prediction. *bioRxiv* **2025**, 637146.
(203) Liu, F.; Xu, H.; Cui, P.; Li, S.; Wang, H.; Wu, Z. NFSA-DTI: A Novel Drug−Target Interaction Prediction Model Using Neural Fingerprint and Self-Attention Mechanism. *Int. J. Mol. Sci.* **2024**, *25*, 11818.
(204) Zeng, X.; Chen, W.; Lei, B. CAT-DTI: cross-attention and Transformer network with domain adaptation for drug-target interaction prediction. *BMC bioinformatics* **2024**, *25*, 141.
(205) Zhang, Q.; Zuo, L.; Ren, Y.; Wang, S.; Wang, W.; Ma, L.; Zhang, J.; Xia, B. FMCA-DTI: A Fragment-oriented method based on a Multihead Cross Attention mechanism to improve Drug-Target Interaction prediction. *Bioinformatics* **2024**, *40*, btae347.
(206) Kang, H.; Goo, S.; Lee, H.; Chae, J.-w.; Yun, H.-y.; Jung, S. Fine-tuning of BERT model to accurately predict drug−target interactions. *Pharmaceutics* **2022**, *14*, 1710.
(207) Zhang, P.-D.; Ma, J.; Chen, T. Escaping the drug-bias trap: using debiasing design to improve interpretability and generalization of drug-target interaction prediction. *bioRxiv* **2024**, 612771.
(208) Qian, Y.; Li, X.; Wu, J.; Zhang, Q. MCL-DTI: using drug multimodal information and bi-directional cross-attention learning method for predicting drug−target interaction. *BMC bioinformatics* **2023**, *24*, 323.
(209) Wang, J.; He, R.; Wang, X.; Li, H.; Lu, Y. MCF-DTI: Multi-Scale Convolutional Local−Global Feature Fusion for Drug−Target Interaction Prediction. *Molecules* **2025**, *30*, 274.
(210) Li, Z.; Zeng, Y.; Jiang, M.; Wei, B. Deep Drug−Target Binding Affinity Prediction Base on Multiple Feature Extraction and Fusion. *ACS omega* **2025**, *10*, 2020−2032.
(211) Kalemati, M.; Zamani Emani, M.; Koohi, S. InceptionDTA: Predicting Drug-Target Binding Affinity with Biological Context Features and Inception Networks. *Heliyon* **2025**, *11*, No. e42476.
(212) Nguyen, N.-Q.; Park, S.; Gim, M.; Kang, J. MulinforCPI: enhancing precision of compound−protein interaction prediction through novel perspectives on multi-level information integration. *Briefings in Bioinformatics* **2024**, *25*, bbad484.
(213) Zhou, C.; Li, Z.; Song, J.; Xiang, W. TransVAE-DTA: Transformer and variational autoencoder network for drug-target binding affinity prediction. *Computer Methods and Programs in Biomedicine* **2024**, *244*, 108003.
(214) Chen, S.; Li, M.; Semenov, I. MFA-DTI: Drug-target interaction prediction based on multi-feature fusion adopted framework. *Methods* **2024**, *224*, 79−92.
(215) Wang, J.; Xiao, Y.; Shang, X.; Peng, J. Predicting drug−target binding affinity with cross-scale graph contrastive learning. *Briefings in Bioinformatics* **2024**, *25*, bbad516.
(216) Li, Z.; Ren, P.; Yang, H.; Zheng, J.; Bai, F. TEFDTA: a transformer encoder and fingerprint representation combined prediction method for bonded and non-bonded drug−target affinities. *Bioinformatics* **2024**, *40*, btad778.
(217) Zhao, L.; Wang, H.; Shi, S. PocketDTA: an advanced multimodal architecture for enhanced prediction of drug- target affinity from 3D structural data of target binding pockets. *Bioinformatics* **2024**, *40*, btae594.
(218) Sun, J.; Wang, H.; Mi, J.; Wan, J.; Gao, J. MTAF−DTA: multi-type attention fusion network for drug−target affinity prediction. *BMC bioinformatics* **2024**, *25*, 375.
(219) Kumar, R.; Romano, J. D.; Ritchie, M. D. CASTER-DTA: Equivariant Graph Neural Networks for Predicting Drug-Target Affinity. *bioRxiv* **2024**, 625281.
(220) Zhang, Q.; Wei, Y.; Liao, B.; Liu, L.; Zhang, S. MMD-DTA.: A multi-modal deep learning framework for drug-target binding affinity and binding region prediction. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* **2024**, *21*, 2200−2211.
(221) Han, L.; Kang, L.; Guo, Q. Imagedta A simple model for drug−target binding affinity prediction. *ACS omega* **2024**, *9*, 28485−28493.
(222) Voitsitskyi, T.; Stratiichuk, R.; Koleiev, I.; Popryho, L.; Ostrovsky, Z.; Henitsoi, P.; Khropachov, I.; Vozniak, V.; Zhytar, R.; Nechepurenko, D.; et al. 3DProtDTA: a deep learning model for drug-target affinity prediction based on residue-level protein graphs. *RSC Adv.* **2023**, *13*, 10261−10272.
(223) He, H.; Chen, G.; Chen, C. Y.-C. NHGNN-DTA: a node-adaptive hybrid graph neural network for interpretable drug−target binding affinity prediction. *Bioinformatics* **2023**, *39*, btad355.
(224) Yang, Z.; Li, Y.; Matsuda, Y.; Song, L. HMG-DTI: a drug-target interaction prediction framework combining Hierarchical Molecular Graphs and improved Convolutional Block Attention Module. *bioRxiv* **2024**, 628851.
(225) Cleves, A. E.; Jain, A. N. Structure-and ligand-based virtual screening on DUD-E+: performance dependence on approximations to the binding pocket. *J. Chem. Inf. Model.* **2020**, *60*, 4296−4310.
(226) Chen, L.; Cruz, A.; Ramsey, S.; Dickson, C. J.; Duca, J. S.; Hornak, V.; Koes, D. R.; Kurtzman, T. Hidden bias in the DUD-E dataset leads to misleading performance of deep learning in structure-based virtual screening. *PloS one* **2019**, *14*, No. e0220113.
(227) Yazdani-Jahromi, M.; Yousefi, N.; Tayebi, A.; Kolanthai, E.; Neal, C. J.; Seal, S.; Garibay, O. O. AttentionSiteDTI: an interpretable graph-based model for drug-target interaction prediction using NLP sentence-level relation classification. *Briefings in Bioinformatics* **2022**, *23*, bbac272.
(228) Li, J.; Sun, L.; Liu, L.; Li, Z. MIFAM-DTI: a drug-target interactions predicting model based on multi-source information fusion and attention mechanism. *Front. Genet.* **2024**, *15*, 1381997.
(229) Ioannidis, V. N.; Song, X.; Manchanda, S.; Li, M.; Pan, X.; Zheng, D.; Ning, X.; Zeng, X.; Karypis, G. Drkg-drug repurposing knowledge graph for covid-19. *arXiv* **2020**, arXiv:2010.09600.
(230) Xia, X.; Zhu, C.; Zhong, F.; Liu, L. MDTips: a multimodal-data-based drug−target interaction prediction system fusing knowledge, gene expression profile, and structural data. *Bioinformatics* **2023**, *39*, btad411.
(231) Zhang, R.; Hristovski, D.; Schutte, D.; Kastrin, A.; Fiszman, M.; Kilicoglu, H. Drug repurposing for COVID-19 via knowledge graph completion. *J. Biomed. Inf.* **2021**, *115*, 103696.
(232) Wang, Z.; Clark, N. R.; Ma'ayan, A. Drug-induced adverse events prediction with the LINCS L1000 data. *Bioinformatics* **2016**, *32*, 2338−2345.
(233) O'Neil, J.; Benita, Y.; Feldman, I.; Chenard, M.; Roberts, B.; Liu, Y.; Li, J.; Kral, A.; Lejnine, S.; Loboda, A.; et al. An Unbiased Oncology Compound Screen to Identify Novel Combination Strategies. *Molecular cancer therapeutics* **2016**, *15*, 1155−1162.
(234) Rafiei, F.; Zeraati, H.; Abbasi, K.; Ghasemi, J. B.; Parsaeian, M.; Masoudi-Nejad, A. DeepTraSynergy: drug combinations using multimodal deep learning with transformers. *Bioinformatics* **2023**, *39*, btad438.
(235) Ortiz, M. A.; Noble, D.; Sorokin, E. P.; Kimble, J. A new dataset of spermatogenic vs. oogenic transcriptomes in the nematode Caenorhabditis elegans. *G3: Genes, Genomes, Genetics* **2014**, *4*, 1765−1772.
(236) Wang, W.; Yang, S.; Zhang, X.; Li, J. Drug repositioning by integrating target information through a heterogeneous network model. *Bioinformatics* **2014**, *30*, 2923−2930.
(237) Olayan, R. S.; Ashoor, H.; Bajic, V. B. DDR: efficient computational method to predict drug−target interactions using

https://doi.org/10.1021/acsomega.5c04997
ACS Omega 2025, 10, 61048−61075

graph mining and machine learning approaches. *Bioinformatics* **2018**, *34*, 1164−1173.
(238) Yamanishi, Y.; Araki, M.; Gutteridge, A.; Honda, W.; Kanehisa, M. Prediction of drug−target interaction networks from the integration of chemical and genomic spaces. *Bioinformatics* **2008**, *24*, i232−i240.
(239) Zheng, Y.; Peng, H.; Zhang, X.; Gao, X.; Li, J. Predicting drug targets from heterogeneous spaces using anchor graph hashing and ensemble learning. In *2018 International Joint Conference on Neural Networks (IJCNN)*, 2018; pp 1−7.
(240) Chandak, P.; Huang, K.; Zitnik, M. Building a knowledge graph to enable precision medicine. *Scientific Data* **2023**, *10*, 67.
(241) Wang, T.; Deng, Q.; Hu, J.; Wang, Y.; Peng, J.; Chen, J.; Shang, X. A multi-scale feature fusion model based on biological knowledge graph and transformer-encoder for drug-drug interaction prediction. *bioRxiv* **2024**, 575305.
(242) Halloran, M. O.; Boyle, C.; Kehoe, B.; Bagkeris, E.; Mallon, P.; Post, F. A.; Vera, J.; Williams, I.; Anderson, J.; Winston, A.; et al. Polypharmacy and Drug−drug Interactions in Older and Younger People Living with HIV: The POPPY Study. *Antiviral therapy* **2019**, *24*, 193−201.
(243) Pham, T.; Ghafoor, M.; Grañana-Castillo, S.; Marzolini, C.; Gibbons, S.; Khoo, S.; Chiong, J.; Wang, D.; Siccardi, M. DeepARV: ensemble deep learning to predict drug-drug interaction of clinical relevance with antiretroviral therapy. *NPJ. systems biology and applications* **2024**, *10*, 48.
(244) Isert, C.; Atz, K.; Jiménez-Luna, J.; Schneider, G. QMugs, quantum mechanical properties of drug-like molecules. *Scientific Data* **2022**, *9*, 273.
(245) Su, M.; Yang, Q.; Du, Y.; Feng, G.; Liu, Z.; Li, Y.; Wang, R. Comparative assessment of scoring functions: the CASF-2016 update. *J. Chem. Inf. Model.* **2019**, *59*, 895−913.
(246) Miranda, A.; Mehryary, F.; Luoma, J.; Pyysalo, S.; Valencia, A.; Krallinger, M. Overview of DrugProt BioCreative VII track: quality evaluation and large scale text mining of drug-gene/protein relations. In *Proceedings of the seventh BioCreative challenge evaluation workshop*, 2021; pp 11−21.
(247) Aldahdooh, J.; Tanoli, Z.; Tang, J. Mining drug−target interactions from biomedical literature using chemical and gene descriptions-based ensemble transformer model. *Bioinformatics advances* **2024**, *4*, vbae106.
(248) Metz, J. T.; Johnson, E. F.; Soni, N. B.; Merta, P. J.; Kifle, L.; Hajduk, P. J. Navigating the kinome. *Nat. Chem. Biol.* **2011**, *7*, 200−202.
(249) Tian, X.; Han, P.; Li, X.; Ding, H.; He, W.; Jing, F.; Wang, S.; Wang, X.; Song, T. KGM4DTI: A Depression Knowledge Graph-Driven Multi-semantic Multi-view Computational Framework for Drug-Target Interaction Prediction. In *2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2024; pp 369−376.
(250) Chen, L.; Gu, Y.; Ji, X.; Sun, Z.; Li, H.; Gao, Y.; Huang, Y. Extracting medications and associated adverse drug events using a natural language processing system combining knowledge base and deep learning. *Journal of the American Medical Informatics Association* **2020**, *27*, 56−64.
(251) Magge, A.; Tutubalina, E.; Miftahutdinov, Z.; Alimova, I.; Dirkson, A.; Verberne, S.; Weissenbacher, D.; Gonzalez-Hernandez, G. DeepADEMiner: a deep learning pharmacovigilance pipeline for extraction and normalization of adverse drug event mentions on Twitter. *Journal of the American Medical Informatics Association* **2021**, *28*, 2184−2192.
(252) Nafea, A. A.; Omar, N.; Al-qfail, Z. M. Artificial neural network and latent semantic analysis for adverse drug reaction detection. *Baghdad Science Journal* **2024**, *21*, 226.
(253) Zhou, F.; Uddin, S.. In *Proceedings of the 2024 Australasian Computer Science Week*; Association for Computing Machinery, 2024; pp 65−68.
(254) Li, Y.; Zhao, W.; Dang, B.; Yan, X.; Gao, M.; Wang, W.; Xiao, M. Research on adverse drug reaction prediction model combining knowledge graph embedding and deep learning. In *2024 4th International Conference on Machine Learning and Intelligent Systems Engineering (MLISE)*, 2024; pp 322−329.
(255) Omana, J.; Jeipratha, P.; Devi, K.; Benila, S.; Revathi, K. Personalized drug recommendation system using Wasserstein auto-encoders and adverse drug reaction detection with weighted feed forward neural network (WAES-ADR) in healthcare. *Journal of Informatics and Web Engineering* **2025**, *4*, 332−347.
(256) Johnson, A. E.; Pollard, T. J.; Shen, L.; Lehman, L.-w. H.; Feng, M.; Ghassemi, M.; Moody, B.; Szolovits, P.; Anthony Celi, L.; Mark, R. G. MIMIC-III, a freely accessible critical care database. *Scientific data* **2016**, *3*, 160035.
(257) Lu, Z. PubMed and beyond: a survey of web tools for searching biomedical literature. *Database* **2011**, *2011*, baq036.
(258) Gaulton, A.; Bellis, L. J.; Bento, A. P.; Chambers, J.; Davies, M.; Hersey, A.; Light, Y.; McGlinchey, S.; Michalovich, D.; Al-Lazikani, B.; et al. ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic acids research* **2012**, *40*, D1100−D1107.
(259) Kuhn, M.; Letunic, I.; Jensen, L. J.; Bork, P. The SIDER database of drugs and side effects. *Nucleic acids research* **2016**, *44*, D1075−D1079.
(260) Ryu, J. Y.; Kim, H. U.; Lee, S. Y. Deep learning improves prediction of drug−drug and drug−food interactions. *Proc. Natl. Acad. Sci. U.S.A.* **2018**, *115*, No. E4304.
(261) Saifuddin, K. M.; Bumgardner, B.; Tanvir, F.; Akbas, E. Hygnn: Drug-drug interaction prediction via hypergraph neural network. In *2023 IEEE 39th International Conference on Data Engineering (ICDE)*, 2023; pp 1503−1516.
(262) Su, J.; Qian, Y. DDI-Transform: A neural network for predicting drug-drug interaction events. *Quantitative Biology* **2024**, *12*, 155−163.
(263) Li, X.; Ji, X.; Xu, C.; Hou, J.; Zhao, X.; Peng, G. MFHG-DDI: An enhanced hybrid graph method leveraging multiple features for predicting drug−drug interactions. *IEEE Access* **2024**, *12*, 188424−188434.
(264) Tamir, A.; Yuan, J.-S. KITE-DDI: A Knowledge graph Integrated Transformer Model for accurately predicting Drug-Drug Interaction Events from Drug SMILES and Biomedical Knowledge Graph. *arXiv* **2024**, arXiv:2412.05770.
(265) Wu, D.; Sun, W.; He, Y.; Chen, Z.; Luo, X. Mkg-fenn: A multimodal knowledge graph fused end-to-end neural network for accurate drug−drug interaction prediction. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2024; pp 10216−10224.
(266) Wang, H.; Zhuang, L.; Ding, Y.; Tiwari, P.; Liang, C. EDDINet Enhancing drug−drug interaction prediction via information flow and consensus constrained multi-graph contrastive learning. *Artificial Intelligence in Medicine* **2025**, *159*, 103029.
(267) Zhou, C.; Zhang, X.; Li, J.; Song, J.; Xiang, W. RGDA-DDI: Residual graph attention network and dual-attention based framework for drug-drug interaction prediction. *arXiv* **2024**, arXiv:2408.15310.
(268) Mondal, S.; Datta, D.; Pramanik, S.; Mehra, R. UniGEN-DDI.: Computing drug-drug interactions using a unified graph embedding network. *bioRxiv* **2024**, 615812.
(269) Zhang, Q.; Wei, Y.; Liu, L. A Domain Adaptive Interpretable Substructure-Aware Graph Attention Network for Drug−Drug Interaction Prediction. *Interdisciplinary Sciences: Computational Life Sciences* **2025**, *17*, 773−790.
(270) Sun, J.; Zheng, H. HDN-DDI a novel framework for predicting drug-drug interactions using hierarchical molecular graphs and enhanced dual-view representation learning. *BMC bioinformatics* **2025**, *26*, 28.
(271) Gromova, A. A.; Maida, A. S. ChemBERTaDDI: Transforming Drug-Drug Interaction Prediction with Transformers and Clinical Insights. *bioRxiv* **2025**, 634309.
(272) Tatonetti, N. P.; Ye, P. P.; Daneshjou, R.; Altman, R. B. Data-driven prediction of drug effects and interactions. *Science translational medicine* **2012**, *4*, 125ra31.
(273) Taboureau, O.; Nielsen, S. K.; Audouze, K.; Weinhold, N.; Edsgard, D.; Roque, F. S.; Kouskoumvekaki, I.; Bora, A.; Curpan, R.; Jensen, T. S.; et al. ChemProt: a disease chemical biology database. *Nucleic acids research* **2010**, *39*, D367.

https://doi.org/10.1021/acsomega.5c04997
ACS Omega 2025, 10, 61048−61075

(274) Kanehisa, M. The KEGG database. In *silicosimulation of biological processes: Novartis Foundation Symposium*, 2002; pp 91−103.
(275) Kuhn, M.; von Mering, C.; Campillos, M.; Jensen, L. J.; Bork, P. STITCH: interaction networks of chemicals and proteins. *Nucleic acids research* **2007**, *36*, D684−D688.
(276) Kim, S.; Thiessen, P. A.; Bolton, E. E.; Chen, J.; Fu, G.; Gindulyte, A.; Han, L.; He, J.; He, S.; Shoemaker, B. A.; et al. PubChem Substance and Compound databases. *Nucleic acids research* **2016**, *44*, D1202−D1213.
(277) Ayvaz, S.; Horn, J.; Hassanzadeh, O.; Zhu, Q.; Stan, J.; Tatonetti, N. P.; Vilar, S.; Brochhausen, M.; Samwald, M.; Rastegar-Mojarad, M.; et al. Toward a complete dataset of drug−drug interaction information from publicly available sources. *J. Biomed. Inf.* **2015**, *55*, 206−217.
(278) Zhang, Y.; Feng, Y.; Wu, M.; Deng, Z.; Wang, S. VGAEDTI: drug-target interaction prediction based on variational inference and graph autoencoder. *BMC bioinformatics* **2023**, *24*, 278.
(279) Dehghan, A.; Razzaghi, P.; Abbasi, K.; Gharaghani, S. TripletMultiDTI: multimodal representation learning in drug-target interaction prediction with triplet loss function. *Expert Systems with Applications* **2023**, *232*, 120754.
(280) Toushik Wasi, A.; Hasan Rafi, T.; Islam, R.; Karlo, S.; Chae, D.-K. CADGL: Context-Aware Deep Graph Learning for Predicting Drug-Drug Interactions. *arXiv* **2024**, arXiv:2403.17210.
(281) Cui, Z.; Lin, L.; Chen, Y.; Wang, S.; Zong, Y. Drug Screening and Target Prediction Based on Machine Learning. *World Journal of Innovation and Modern Technology* **2024**, *7*,
(282) Sharma, R.; Saghapour, E.; Chen, J. Y. An NLP-based technique to extract meaningful features from drug SMILES. *Iscience* **2024**, *27*, 109127.
(283) Zhao, X.; Wang, Q.; Zhang, Y.; He, C.; Yin, M.; Zhao, X. CBKG-DTI: Multi-Level Knowledge Distillation and Biomedical Knowledge Graph for Drug-Target Interaction Prediction. *IEEE Journal of Biomedical and Health Informatics* **2025**, *29*, 2284−2296.
(284) Lin, X.; Zhang, X.; Yu, Z.-G.; Long, Y.; Zeng, X.; Yu, P. S. CSCL-DTI: predicting drug-target interaction through cross-view and self-supervised contrastive learning. In *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2024; pp 707−712.
(285) Zeng, X.; Liang, Y.; Yang, P.; Chen, W.; Lei, B. Hypergraph-based Self-supervised Multi-channel Network for Drug-target Interaction Prediction. In *2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2024; pp 395−400.
(286) Lin, X.; Chen, Y.; Yang, W.; Yu, H.; Wu, Z.; Zhang, X. Detecting DTI Using Graph Embedding and Multi-Head Attention Mechanism. In *2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2024; pp 1584−1587.
(287) Szydlik, S.; Taheri, G. DTPPI: predicting drug interactions using a weighted drug-protein network. *bioRxiv* **2025**, 631638.
(288) Zhou, D.; Xu, Z.; Li, W.; Xie, X.; Peng, S. MultiDTI: drug−target interaction prediction based on multi-modal representation learning to bridge the gap between new chemical entities and known heterogeneous network. *Bioinformatics* **2021**, *37*, 4485−4492.
(289) Rappaport, N.; Twik, M.; Plaschkes, I.; Nudel, R.; Iny Stein, T.; Levitt, J.; Gershoni, M.; Morrey, C. P.; Safran, M.; Lancet, D. MalaCards: an amalgamated human disease compendium with diverse clinical and genetic annotation and structured search. *Nucleic acids research* **2017**, *45*, D877−D887.
(290) Keshava Prasad, T.; Goel, R.; Kandasamy, K.; Keerthikumar, S.; Kumar, S.; Mathivanan, S.; Telikicherla, D.; Raju, R.; Shafreen, B.; Venugopal, A.; et al. Human protein reference database2009 update. *Nucleic acids research* **2009**, *37*, D767−D772.
(291) Davis, A. P.; Grondin, C. J.; Johnson, R. J.; Sciaky, D.; Wiegers, J.; Wiegers, T. C.; Mattingly, C. J. Comparative toxicogenomics database (CTD): update 2021. *Nucleic acids research* **2021**, *49*, D1138−D1143.
(292) Desaphy, J.; Bret, G.; Rognan, D.; Kellenberger, E. sc-PDB a 3D-database of ligandable binding sites10 years on. *Nucleic acids research* **2015**, *43*, D399−D404.
(293) Hu, L.; Benson, M. L.; Smith, R. D.; Lerner, M. G.; Carlson, H. A. Binding MOAD (mother of all databases). *Proteins: Struct., Funct., Bioinf.* **2005**, *60*, 333−340.
(294) Wang, R.; Fang, X.; Lu, Y.; Yang, C.-Y.; Wang, S. The PDBbind database: methodologies and updates. *Journal of medicinal chemistry* **2005**, *48*, 4111−4119.
(295) Bourgeas, R.; Basse, M.-J.; Morelli, X.; Roche, P. Atomic analysis of protein-protein interfaces with known inhibitors: the 2P2I database. *PloS one* **2010**, *5*, No. e9598.
(296) Yang, W.; Soares, J.; Greninger, P.; Edelman, E. J.; Lightfoot, H.; Forbes, S.; Bindal, N.; Beare, D.; Smith, J. A.; Thompson, I. R.; et al. Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. *Nucleic acids research* **2012**, *41*, D955−D961.
(297) Sondka, Z.; Dhir, N. B.; Carvalho-Silva, D.; Jupe, S.; Madhumita; McLaren, K.; Starkey, M.; Ward, S.; Wilding, J.; Ahmed, M.; et al. COSMIC: a curated database of somatic variants and clinical data for cancer. *Nucleic Acids Res.* **2024**, *52*, D1210−D1217.
(298) Munk, C.; Isberg, V.; Mordalski, S.; Harpsøe, K.; Rataj, K.; Hauser, A.; Kolb, P.; Bojarski, A.; Vriend, G.; Gloriam, D. GPCRdb: the G protein-coupled receptor database−an introduction. *Br. J. Pharmacol.* **2016**, *173*, 2195−2207.
(299) Hewett, M.; Oliver, D. E.; Rubin, D. L.; Easton, K. L.; Stuart, J. M.; Altman, R. B.; Klein, T. E. PharmGKB: the pharmacogenetics knowledge base. *Nucleic acids research* **2002**, *30*, 163−165.
(300) Rezaei, Z.; Ebrahimpour-Komleh, H.; Eslami, B.; Chavoshinejad, R.; Totonchi, M. Adverse drug reaction detection in social media by deep learning methods. *Cell Journal* **2020**, *22*, 319.
(301) Nguyen, D. P.; Bao Ho, T. Drug-Drug Interaction Extraction from Biomedical Texts via Relation BERT. In *2020 RIVF International Conference on Computing and Communication Technologies (RIVF)*, 2020; pp 1−7.
(302) Xie, Z.; Zhu, R.; Zhang, M.; Liu, J. SparseMult: A Tensor Decomposition model based on Sparse Relation Matrix. In *2022 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)*, 2022; pp 761−764.
(303) Nickel, M.; Rosasco, L.; Poggio, T. Holographic embeddings of knowledge graphs. In *Proceedings of the AAAI conference on artificial intelligence*, 2016.
(304) Kuzin, M.; Haen, E.; Kuzo, N.; Endres, K.; Hiemke, C.; Paulzen, M.; Schoretsanitis, G. Assessing pharmacokinetic correlates of escitalopram-related adverse drug reactions. *Therapeutic Drug Monitoring* **2024**, *46*, 246−251.
(305) Masumshah, R.; Aghdam, R.; Eslahchi, C. A neural network-based method for polypharmacy side effects prediction. *BMC bioinformatics* **2021**, *22*, 385.
(306) Gao, M.; Chen, L.; He, X.; Zhou, A. Bine.: Bipartite network embedding. In *41st international ACM SIGIR conference on research & development in information retrieval*, 2018; pp 715−724.
(307) Forster, D. T.; Li, S. C.; Yashiroda, Y.; Yoshimura, M.; Li, Z.; Isuhuaylas, L. A. V.; Itto-Nakama, K.; Yamanaka, D.; Ohya, Y.; Osada, H.; et al. BIONIC: biological network integration using convolutions. *Nat. Methods* **2022**, *19*, 1250−1261.
(308) Feng, Z.-H.; Kittler, J.; Awais, M.; Wu, X.-J. Rectified wing loss for efficient and robust facial landmark localisation with convolutional neural networks. *International Journal of Computer Vision* **2020**, *128*, 2126−2145.
(309) Zhang, B.; Niu, D.; Zhang, L.; Zhang, Q.; Li, Z. MSH-DTI: multi-graph convolution with self-supervised embedding and heterogeneous aggregation for drug-target interaction prediction. *BMC bioinformatics* **2024**, *25*, 275.
(310) Zhao, W.; Yu, Y.; Liu, G.; Liang, Y.; Xu, D.; Feng, X.; Guan, R. MSI-DTI: predicting drug-target interaction based on multi-source information and multi-head self-attention. *Briefings in Bioinformatics* **2024**, *25*, bbae238.
(311) Benesty, J.; Chen, J.; Huang, Y.; Cohen, I.; Benesty, J.; Chen, J.; Huang, Y.; Cohen, I. Pearson correlation coefficient. In *Noise reduction in speech processing*; Springer, 2009; pp 1−4..

https://doi.org/10.1021/acsomega.5c04997
ACS Omega 2025, 10, 61048−61075

(312) De Winter, J. C.; Gosling, S. D.; Potter, J. Comparing the Pearson and Spearman correlation coefficients across distributions and sample sizes: A tutorial using simulations and empirical data. *Psychological methods* **2016**, *21*, 273.
(313) Cha, Y.; Erez, T.; Reynolds, I. J.; Kumar, D.; Ross, J.; Koytiger, G.; Kusko, R.; Zeskind, B.; Risso, S.; Kagan, E.; et al. Drug Repurposing from the Perspective of Pharmaceutical Companies. *Br. J. Pharmacol.* **2018**, *175*, 168−180.
(314) Breckenridge, A.; Jacob, R. Overcoming the Legal and Regulatory Barriers to Drug Repurposing. *Nat. Rev. Drug Discovery* **2019**, *18*, 1−2.
(315) Talevi, A.; Bellera, C. L. Challenges and Opportunities with Drug Repurposing: Finding Strategies to Find Alternative Uses of Therapeutics. *Expert Opinion on Drug Discovery* **2020**, *15*, 397−401.
(316) Wang, T.; Yang, J.; Xiao, Y.; Wang, J.; Wang, Y.; Zeng, X.; Wang, Y.; Peng, J. DFinder: a novel end-to-end graph embedding-based method to identify drug−food interactions. *Bioinformatics* **2023**, *39*, btac837.
(317) Jarab, A. S.; Al-Qerem, W.; Alajlouni, H. Y.; Alzoubi, K. H.; Abu Heshmeh, S.; Mukattash, T. L.; Naser, A. Y.; Al Hamarneh, Y. N. Public knowledge and attitude towards drug-food interactions: implications for improved public health safety. *International Journal of Environmental Health Research* **2024**, *34*, 3025−3035.
(318) Alhubail, S. A.; Alharthi, M. M.; Alsayyah, F. F.; Younis, N. S. Healthcare professionals and undergraduate students knowledge toward drug-food interactions in the eastern region of Saudi Arabia. *Journal of Multidisciplinary Healthcare* **2023**, *Volume 16*, 2883−2892.
(319) Creswell, A.; White, T.; Dumoulin, V.; Arulkumaran, K.; Sengupta, B.; Bharath, A. A. Generative Adversarial Networks: An Overview. *IEEE Signal Processing Magazine* **2018**, *35*, 53−65.
(320) McAfee, R. P.; Reny, P. J. Correlated Information and Mechanism Design. *Econometrica: Journal of the Econometric Society* **1992**, *60*, 395−421.

https://doi.org/10.1021/acsomega.5c04997
ACS Omega 2025, 10, 61048−61075