

Received 14 January 2025, accepted 13 February 2025, date of publication 28 February 2025, date of current version 12 March 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3546925

METHODS

Evaluating Critical Success Factors in AI-Driven Drug Discovery Using AHP: A Strategic Framework for Optimization

AMIR MOHAMED TALIB¹, AREEN METIB AL-HGAISH^{2,3}, RODZIAH BINTI ATAN³,
ABDULAZIZ ALSHAMMARI¹, FAHAD OMAR ALOMARY¹,
RAZALI YAAKOB³, (Member, IEEE), ABDULAZIZ ALSAHLI¹,
AND MOHD HAFEEZ OSMAN³

¹College of Computer and Information Sciences, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh 11564, Saudi Arabia

²Electrical Engineering Department, Al-Zaytoonah University of Jordan, Amman 11733, Jordan

³Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, Serdang 43400, Malaysia

Corresponding author: Areen Metib Al-Hgaish (areen.metib@gmail.com)

This work was supported by the Deanship of Scientific Research at Imam Mohammad Ibn Saud Islamic University under Grant RG-23021.

ABSTRACT Artificial Intelligence (AI) is reshaping drug discovery by accelerating the identification of therapeutic candidates and reducing development timelines and costs. However, its effectiveness depends on addressing key success factors that influence AI integration. This study presents a novel framework using the Analytic Hierarchy Process (AHP) to systematically evaluate and rank these factors, addressing a crucial gap in strategic planning for AI adoption in pharmaceutical research. The framework comprises six key criteria: Data Quality and Management (DQM), Algorithm Performance and Optimization (APO), Interpretability and Explainability (IE), Regulatory Compliance and Ethical Considerations (RCEC), Computational Efficiency and Scalability (CES), and Validation and Experimental Confirmation (VEC). Expert-driven pairwise comparisons identified Accuracy (ACC), Generalizability (GEN), and Experimental Validation (EV) as top priorities, highlighting the importance of reliable data, robust algorithms, and rigorous validation processes to ensure trustworthy AI outputs. This research contributes to strategic AI adoption by addressing data inconsistencies, algorithmic bias, and scalability limitations. The proposed framework enhances AI applications' efficiency, scalability, and ethical alignment, promoting the development of transparent and reliable drug discovery systems. This comprehensive evaluation is a valuable resource for researchers and industry professionals, facilitating the strategic adoption of AI and bridging the gap between computational predictions and real-world therapeutic outcomes.

INDEX TERMS Drug discovery, AI, AI-driven, analytic hierarchy process, critical factors, decision-making, model performance.

I. INTRODUCTION

Artificial Intelligence (AI) is transforming the drug discovery process by speeding up the development of new therapeutics and reducing high costs, extended timelines, and high failure rates inherent in traditional pharmaceutical innovation [1], [2]. The successful integration of AI-driven approaches in

pharmaceutical research relies on several critical factors. High-quality data availability and preprocessing remain fundamental to AI effectiveness, as inconsistencies in biomedical datasets can significantly impact model performance and reproducibility [3].

Machine learning (ML) applications enhance all stages of drug discovery by enabling AI to support data-driven decision-making and boost clinical trial success rates. However, challenges remain, including the interpretability of ML

The associate editor coordinating the review of this manuscript and approving it for publication was Catherine Fang.

models and the necessity for systematic, high-dimensional biological datasets, which can hinder AI's reliability in real-world applications [4].

Deep learning (DL)-based approaches have demonstrated remarkable advancements in predicting drug-target interactions (DTIs), virtual screening, and de novo drug design [5]. By leveraging neural networks and molecular representations, AI can accelerate lead optimization while enhancing the precision of drug selection. However, the integration of AI in drug discovery is still hindered by data biases, generalization limitations, and regulatory constraints [6]. Addressing these challenges requires interdisciplinary collaboration and improved data-sharing frameworks to maximize AI's impact on pharmaceutical innovation. Additionally, regulatory concerns surrounding AI-driven drug development are becoming increasingly significant as evolving policies struggle to keep pace with technological advancements [7]. Given that traditional drug development timelines exceed 12 years and cost more than \$2.5 billion per drug [8], AI-driven innovations offer a compelling strategy to optimize drug discovery, reduce costs, and improve approval rates.

AI models have demonstrated their potential to optimize key stages of drug development, such as target identification, hit generation, and chemical optimization [9], [10]. ML and DL methodologies have shown remarkable promise in preclinical research by leveraging computational modeling to predict outcomes and discover novel compounds [11], [12]. Despite these advancements, challenges such as data scarcity, heterogeneity, and class imbalance impede the practical deployment of AI models in real-world pharmaceutical applications [13], [14]. Moreover, many advanced AI systems function as opaque "black boxes," making it difficult for healthcare professionals and regulatory agencies to fully trust their predictions without enhanced model interpretability [14], [15].

The ethical application of AI in drug discovery introduces further complexities, such as ensuring algorithmic fairness, protecting data privacy, and maintaining regulatory compliance to guarantee accountability and reliability [16], [17]. Aligning AI-generated predictions with experimental validation is a significant challenge, as discrepancies between computational outcomes and clinical trial results can undermine the reliability of AI-driven drug candidates [18]. Industrial-scale AI-powered platforms have sought to integrate multi-omics, ML, and computational precision medicine to refine target selection and reduce the risk of clinical failure, yet challenges remain in standardizing data integration and model interpretability [18]. ML-based drug repositioning methods also leverage large-scale datasets and predictive modeling to identify novel therapeutic applications for existing drugs. Still, the lack of interpretability in model decisions continues to be a limiting factor [19]. The quality, completeness, and accessibility of bioactivity data further influence the reliability of AI-based drug discovery, as inconsistency in experimental assays and deposition

protocols hinders reproducibility and cross-validation [20]. While techniques such as transfer learning and explainable AI (XAI) have been introduced to enhance model adaptability to novel drug targets [21], [22], the absence of standardized benchmarks and incomplete bioassay datasets continue to restrict the scalability and generalization of AI models across diverse disease areas [5], [19].

To address these multifaceted challenges, this study presents a strategic framework that utilizes the AHP to identify and prioritize essential success criteria in AI-driven drug discovery. The proposed framework is driven by expert insights, offering a scientific approach to optimize resources and enhance the effectiveness of AI-driven drug discovery. Unlike previous studies focusing on isolated computational advancements, this research systematically integrates expert-driven prioritization to evaluate the most critical factors influencing AI-driven drug discovery. By applying a structured decision-making approach, this study ensures that AI methodologies align with real-world pharmaceutical constraints, strengthening their scientific validity and industry relevance. This contribution enhances resource allocation and supports the development of AI models that meet technical and regulatory requirements, ultimately improving the practical applicability of AI-driven drug discovery frameworks. It assesses key challenges and limitations highlighted in previous studies by prioritizing critical criteria identified from the literature. This comprehensive approach aims to alleviate bottlenecks in the drug development pipeline, enhance predictive performance, improve model interpretability, and ensure regulatory compliance. Ultimately, the framework seeks to accelerate drug discovery processes, improve approval rates for new therapeutic candidates, and streamline the path toward innovative treatments.

This paper contributes to the growing body of research on strategic planning for AI-driven drug discovery by:

- **Bridging the Gap Between AI and Pharmaceutical Sciences:** This research establishes a scientific framework that aligns AI-driven strategies with the practical needs of drug discovery. By integrating expert opinions from machine learning, deep learning, and AI applications, the study enhances the applicability and robustness of AI in pharmaceutical research.
- **Providing a Structured Decision-Making Framework:** The study enhances the scientific community by offering a structured decision-making framework for AI-driven drug discovery. Utilizing an AHP-based prioritization process, it systematically evaluates and ranks critical technical, regulatory, and operational factors. This approach helps researchers and industry professionals navigate the complexities of integrating AI into drug development.
- **This study Advancing Multi-Criteria Analysis in AI Implementation:** This study enriches existing literature by introducing a comprehensive multi-criteria analysis that encompasses six critical areas:

1. Data Quality Management (DQM).
2. Algorithm Performance and Optimization (APO).
3. Interpretability and Explainability (IE).
4. Regulatory Compliance and Ethical Considerations (RCEC).
5. Computational Efficiency and Scalability (CES).
6. Validation and Experimental Confirmation (VEC).

By identifying 24 interrelated success factors within these categories, the research offers a more holistic understanding of the challenges in AI-driven drug discovery. It is a reference for future studies in computational drug development.

- **Enhancing Transparency and Reliability in AI-Driven Drug Discovery:** A significant challenge in AI adoption for drug discovery is ensuring transparency, interpretability, and regulatory compliance. This study directly highlights this issue by structuring expert-driven prioritization to enhance the reliability of AI applications in pharmaceutical research. The proposed framework ensures that AI adoption is based on scientifically validated criteria rather than ad hoc or isolated technical advancements.
- **Facilitating Efficient Resource Allocation and Strategic Planning:** This study offers pharmaceutical companies and regulatory bodies a strategic roadmap to optimize resource allocation. By employing AHP for pairwise comparisons, decision-makers can prioritize factors that most significantly impact the success of AI adoption in drug discovery. This structured approach promotes efficiency and cost-effectiveness in pharmaceutical research and development.
- **Laying a Foundation for Future Research and Policy Development:** This research provides a foundation for future advancements in AI-driven drug discovery, allowing policymakers, researchers, and industry leaders to build on a scientific framework. It serves as a benchmark for evaluating AI strategies in pharmaceutical sciences and paves the way for further refinement through empirical validation and real-world implementation.

The AHP-based framework fills a crucial research gap by creating a comprehensive, expert-driven system that systematically evaluates and prioritizes key factors for AI-driven drug discovery. It transcends technical discussions to offer an integrative approach encompassing scientific, regulatory, and operational considerations. This methodology ensures a balanced and validated AI adoption strategy, significantly advancing theoretical and practical aspects of modern pharmaceutical development.

Many experts with diverse research interests were invited to ensure the AHP framework reflects a well-rounded and practical understanding of the challenges in AI-driven drug discovery. This diversity encompasses specialists in ML, DL, and other computational methods, all of whom have significantly contributed to drug discovery through their research and publications. The selection process prioritized academic achievements, professional experience, and active

involvement in developing computational models, ensuring a broad spectrum of perspectives and practical expertise was captured.

From the invited pool, 21 experts completed the AHP questionnaire, representing the variety of research interests necessary to address the multifaceted challenges in this domain. The chosen number of experts aligns with the AHP methodology, which relies on qualitative, pairwise comparisons rather than large statistical samples to derive meaningful results [23]. This approach ensures critical factors are prioritized accurately based on expert knowledge and experience. The participants' diverse backgrounds enhance the proposed framework's robustness and applicability, fostering innovative solutions to optimize AI-driven drug discovery.

Table 1 presents a comprehensive summary of 21 experts, emphasizing their contributions and research focuses on computational drug discovery. This table shows the diverse expertise involved in the study, which spans a wide range of topics. Integrating various research areas strengthens the robustness of the AHP framework used in the research.

The structure of this study is outlined as follows: Section II provides an in-depth literature review on key research areas and methodologies related to AI-driven drug discovery; Section III outlines the study's methodology, detailing the approach and analysis steps taken; Section IV presents and discusses the results obtained; Section V highlights the implications, limitations, and future research directions stemming from the findings; Section VI concludes the paper with a summary of contributions and insights.

II. LITERATURE REVIEW

This section defines and analyzes the current state of AI-driven methods in drug discovery, addressing challenges, limitations of previous studies, diverse applications of AHP, and its potential to improve drug discovery efforts.

A. DRUG DISCOVERY

Drug The discovery and development of a new drug is a complex and expensive process, costing up to \$2 billion for a single drug [38]. This resource-intensive journey consists of several phases: target identification and validation, hit identification, lead optimization, preclinical testing, clinical trials, and regulatory approval (Fig.1) [39]. On average, drug discovery takes over 12 years, with a failure rate exceeding 90% [8].

Drug discovery typically begins with medicinal chemists generating a library of lead compounds, which are assessed through structure-activity relationships (SARs) to evaluate their in vitro efficacy and preclinical in vivo safety. Subsequently, promising candidates are subjected to formulation, stability assessments, scale-up production, and chronic safety studies in animal models before advancing to clinical trials [40], [41]. Despite these rigorous efforts, many drug

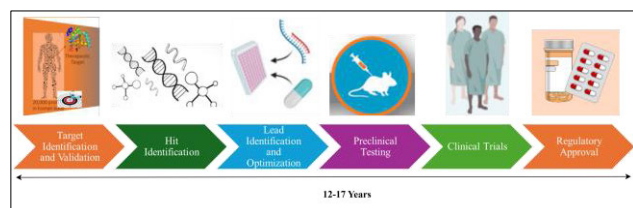
TABLE 1. Summary of expert contributions to AI-driven drug discovery research.

Expert	Research Title	Key Focus Area(s)	Discussion Issues of the Article
1	Comprehensive Ensemble in QSAR Prediction for Drug Discovery [24]	Ensemble learning, integration of diverse models, and meta-learning for drug prediction.	The study addresses data integration challenges by incorporating diverse data representations like fingerprints and SMILES to enhance predictive accuracy. It employs ensemble learning techniques to improve generalizability, mitigate overfitting, and ensure robust performance in ML frameworks.
2	The Changing Scenario of Drug Discovery Using AI to Deep Learning [25]	AI-driven drug discovery advancements include target identification, ADMET prediction, de novo drug design, and toxicity assessment.	This research addresses data curation and biological dataset integration to enhance the generalizability of ML/DL models. It promotes explainable AI to overcome the limitations of black-box models and highlights the importance of experimental validation for the reliability of AI-predicted drug candidates.
3, 4	Privacy-Preserving Techniques for Decentralized and Secure Machine Learning in Drug Discovery [26]	Implement secure ML frameworks, including federated learning, homomorphic encryption, and differential privacy for drug discovery applications.	The study highlights secure data integration and privacy-preserving techniques such as federated learning and differential privacy to maintain data confidentiality and scalability. It emphasizes the importance of decentralized learning frameworks, benchmarking, and validation for robust and ethical ML applications in drug discovery.
5	Machine Learning in Drug Discovery: A Critical Review of Applications and Challenges [6]	ML applications in drug discovery include target identification, hit discovery, lead optimization, and predictive toxicology.	The study emphasizes that robust data integration and FAIR principles address challenges such as data noise, biases, and availability. It emphasizes explainable AI for model interpretability and stresses the need for experimental validation, benchmarking, and regulatory alignment to ensure reliability and adoption.
6	GeoT: A Geometry-Aware Transformer for Reliable Molecular Property Prediction [27]	Incorporation of molecular geometric information in transformer models; enhanced interpretability and prediction reliability.	The paper introduces a geometry-aware transformer model that enhances predictive accuracy and interpretability by connecting molecular structures, ensuring completeness and accuracy to outcomes. It optimizes resource usage and computational scalability, ensuring reliability in large-scale applications.
7, 8, 9	Deep Learning Pipeline for Accelerating Virtual Screening in Drug Discovery [28]	Development of VirtuDockDL, an AI-based tool integrating DL, molecular docking, and virtual screening for drug discovery.	The study employs advanced preprocessing techniques using SMILES and descriptors to enhance data accuracy. Graph Neural Networks are leveraged for superior predictive performance, addressing overfitting challenges. Validation methods are benchmarked against traditional approaches and achieve high accuracy in identifying active drug candidates, demonstrating robustness and reliability.
10	Using Artificial Intelligence and Machine Learning Approaches to Enhance Cancer Therapy And Drug Discovery [29]	Application of AI and ML in drug discovery, focusing on multitarget therapies, natural compounds, and anticancer drugs.	The research focuses on data integration and FAIR principles to manage diverse datasets in cancer drug discovery. AI-driven methods are utilized for drug target prediction and virtual screening, while experimental validation ensures the accuracy and reliability of identified anticancer compounds.
11, 12	A Recruitment Big Data Approach to Interplay of the Target Drugs [30]	Use of big data frameworks to analyze and preprocess drug datasets; clustering techniques and feature selection for optimization.	The study utilizes big data frameworks and clustering techniques to address missing data and optimize dataset structure. Preprocessing enhances data stability, benchmarking confirms reliability and experimental validation reinforces robust outcomes.
13	Machine Learning Applications in Drug Discovery [31]	Integrating ML techniques like neural networks, reinforcement learning, and QSAR modeling for drug discovery.	The paper emphasizes accurately handling large datasets while addressing overfitting challenges. Generalizable ML models are developed for reliable toxicity predictions and drug-target interaction validation, ensuring robust and effective outcomes.
14	Novel Numerical Simulation of Drug Solubility in Supercritical CO ₂ Using Machine Learning Technique [32]	Application of ML techniques like Kernel Ridge Regression (KRR), Least Angle Regression (LAR), and Multilayer Perceptron (MLP) to optimize drug solubility in supercritical CO ₂ .	The study leverages ML models to predict and optimize drug solubility with high accuracy (R-squared scores above 0.99 for the MLP model). It addresses challenges in data modeling and ensures robust computational efficiency through advanced regression methods and neural networks.
15, 16	Predictive Modeling of Skin Permeability for Molecules Using AI Algorithms [33]	Application of advanced AI models (LGBM, XGBoost, Gradient Boosting) for predicting skin drugs.	Advanced AI models such as LGBM and XGBoost are utilized to achieve high predictive accuracy. Descriptor-based clustering enhances data integrity, and experimental validation ensures the practical applicability of predictions in drug discovery.
17	Machine Learning for Drug Discovery: Bridging Computational Science and Medicine [34]	ML applications in drug discovery, focusing on target identification, drug repurposing, predictive toxicology, and clinical trial optimization.	The study ensures data accuracy and reliability using structured datasets and comprehensive preprocessing. Through case studies and real-world applications, it showcases improved predictive accuracy and target validation, enhancing drug discovery.
18	Chemocentric Informatics Approach to Drug Discovery: Identification and Experimental Validation of Selective Estrogen Receptor Modulators as Ligands of 5-Hydroxytryptamine-6 Receptors and as Potential Cognition Enhancers [35]	Integrative cheminformatics, QSAR modeling, and gene-expression profiling for Alzheimer's drug discovery.	The study presents a comprehensive informatics framework integrating QSAR-based virtual screening with gene signature network mining. This method identifies potential ligands for 5-HT ₆ receptors and verifies their activity through experimental assays. It addresses challenges related to data integration, accuracy, and cross-validation of computational predictions. The study highlights the significance of multitarget modeling and experimental validation in enhancing drug discovery success rates.

TABLE 1. (Continued.) Summary of expert contributions to AI-driven drug discovery research.

19	Identification of Non-Classical hCA XII Inhibitors Using Computational Drug Design Approaches [36]	Employ computational methods such as pharmacophore modeling, molecular docking, and dynamics to discover new hCA XII inhibitors.	The study focuses on computational drug design methods, addressing isoform specificity, side effects, and resistance in traditional inhibitors. It highlights using ligand- and structure-based pharmacophore models to discover non-classical inhibitors with significant therapeutic potential.
20, 21	Predicting Medical Drug Towards Meaningful Data Using Machine Learning [37]	ML approaches for drug classification and alternative drug identification using structured datasets.	This research focuses on advanced ML classifiers like Random Forest, Naive Bayes, and Decision Tree for robust drug classification and alternative identification. Preprocessing and validation methods ensure reliability and accuracy in predictions.

candidates fail to reach the market due to formulation challenges, even when demonstrating potent biological activity. Figure 1 illustrates the drug development processes.

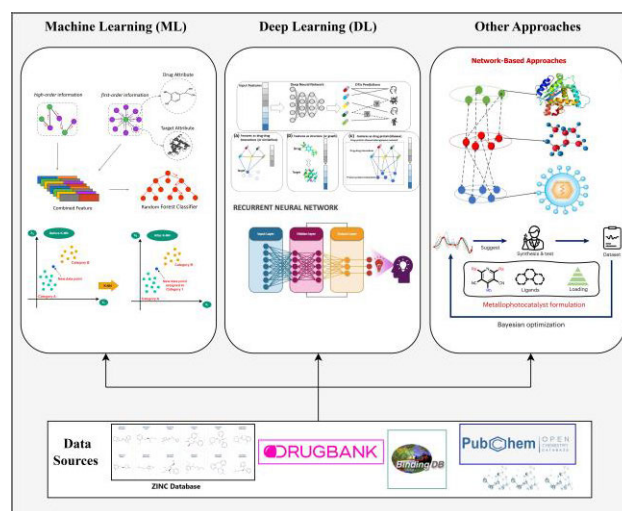
**FIGURE 1. Overview of the drug discovery process.**

Recent AI advancements promise to enhance drug discovery by using ML and DL techniques to expedite hit identification and lead optimization, which are critical early phases of drug development. These technologies enable rapid screening of millions of compounds, providing predictive insights into binding affinities and pharmacological properties, thus improving the precision and efficiency of the selection process [42]. Furthermore, AI-based models play a pivotal role in refining the lead optimization phase by predicting drug-likeness and potential toxicities, thereby improving the probability of selecting viable candidates. Despite the significant potential of AI-driven approaches, challenges remain, including the need for high-quality datasets, regulatory barriers, and the requirement for experimental validation to confirm computational predictions [16], [19]. Overcoming these limitations is crucial for effectively integrating AI into the pharmaceutical industry's workflows and expediting the delivery of safe and effective therapeutics to market.

B. COMPUTATIONAL DRUG DISCOVERY APPROACHES

Computational methods such as ML and DL have emerged as indispensable tools in modern drug discovery. These advanced techniques enable the comprehensive analysis of vast amounts of biological data, molecular interactions, and chemical structures with remarkable efficiency and accuracy [2]. Molecular docking, virtual screening, and quantitative structure-activity relationship (QSAR) modeling are widely used to predict drug efficacy and safety profiles [43]. Public databases like ZINC, BindingDB, and PubChem provide valuable datasets that enable computational techniques to predict interactions between drugs and biological targets.

These databases give access to vast chemical and biological data repositories essential for in-silico drug discovery [44], [45]. However, heterogeneous data sources pose a significant challenge due to the varying formats and quality of available datasets, making it difficult to integrate and analyze them seamlessly [13], [46]. Moreover, data warehousing tools such as SWISS-PROT, BIOMOLQUEST, and others have been developed to address data integration issues, yet achieving data standardization and curation remains critical [47].

**FIGURE 2. AI-driven drug discovery techniques.**

In this context, Figure 2 provides an overview of computational approaches categorized into three main groups: ML, DL, and other methods. It includes examples of popular techniques such as random forest classification, k-nearest neighbors (k-NN), recurrent neural networks (RNNs), network-based methods, and Bayesian optimization strategies.

The figure illustrates how public databases like ZINC and PubChem [44], [45]. Supply these models with extensive chemical and biological data, enhancing training and prediction reliability. Researchers can address heterogeneity by integrating data from multiple sources and utilizing data warehousing tools, ultimately accelerating drug discovery through AI-driven solutions. The following subsections will explore these approaches.

1) MACHINE LEARNING (ML)

ML is now crucial in computational drug discovery. By analyzing complex biological datasets, ML techniques provide powerful predictive abilities. These techniques significantly enhance drug-target interaction (DTI) prediction, virtual screening, and lead compound identification [14]. Popular algorithms such as Random Forest, Support Vector Machines (SVM), and k-nearest Neighbors (kNN) identify patterns within large-scale, high-dimensional biological data, streamlining the drug discovery process [13]. Despite their effectiveness, these models often struggle when trained on small or imbalanced datasets. They face challenges such as overfitting, which limits their generalizability to new data [6], [48].

Several strategies have been implemented to improve model robustness and address these limitations. Data augmentation and resampling techniques balance class distributions, while cost-sensitive learning introduces penalties for misclassifications, enhancing the performance of ML models in drug discovery applications [48], [49]. By incorporating these advancements, ML-based frameworks can provide more accurate and reliable predictions, contributing to more efficient and targeted drug development pipelines.

2) DEEP LEARNING (DL)

DL represents a significant breakthrough in AI, fundamentally transforming drug discovery through its advanced modeling capabilities that effectively learn intricate and non-linear patterns from extensive datasets [50]. These techniques empower researchers to process and interpret complex biological data, driving more effective and innovative drug development approaches. A prominent innovation in DL is the application of graph-based neural networks, which enhance prediction accuracy by representing molecular structures as graphs [5].

Furthermore, DL architectures such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) have demonstrated exceptional performance in predicting drug-target interactions, accelerating drug repurposing efforts, and optimizing molecular properties [51], [52].

However, with these advancements, DL models often face the challenge of limited interpretability, commonly called the black-box problem, where the decision-making processes remain opaque, hindering acceptance in clinical and regulatory environments [53]. To improve transparency, Explainable AI (XAI) approaches such as Local Interpretable Model-Agnostic Explanations (LIME) have been introduced. These offer insights into model predictions and foster trust in AI-driven outputs [21]. Addressing the interpretability issue can help DL frameworks achieve broader adoption and integration in real-world drug discovery workflows, paving the way for safer and more efficient pharmaceutical innovations.

3) OTHER APPROACHES USED AI-DRIVEN DRUG DISCOVERY

Building upon the foundations of ML and DL, various complementary methodologies have emerged, playing critical

roles in advancing drug discovery processes. Network-based approaches, for instance, enable the identification of novel drug-disease associations by analyzing complex relationships among drugs, diseases, and biological targets. This network-centric analysis aids in uncovering previously unrecognized connections, thereby bolstering drug repurposing efforts and expediting the identification of new therapeutic applications for existing drugs [54].

Transfer learning is pivotal, especially when limited labeled data are available. By leveraging models pre-trained on extensive datasets, transfer learning facilitates fine-tuning these models for specific tasks within drug discovery, thereby enhancing efficiency and precision. This approach enables researchers to overcome data scarcity issues by transferring relevant knowledge across related domains, ultimately accelerating hypothesis generation and validation [55]. Moreover, multitasking learning further enhances learning outcomes by training models across multiple related tasks, which helps mitigate data scarcity issues. Additionally, Bayesian optimization and ensemble learning techniques are often utilized to boost model performance and tackle overfitting, enhancing scalability and robustness in AI-driven drug discovery processes [56], [57]. Ontology-based integration tools, like the Ontology Web Language and RDF Schema, address data heterogeneity by integrating biomedical data from diverse sources. However, scalability remains challenging with large datasets of varying formats [58], [59].

Explainable AI (XAI) techniques, such as Local Interpretable Model-agnostic Explanations (LIME), have emerged as crucial tools for increasing transparency, fostering model acceptance, and addressing ethical and regulatory requirements in clinical applications [60], [61]. Furthermore, big data analytics play a significant role in managing and analyzing diverse datasets, supporting pipelines that integrate data from major public databases like ZINC, BindingDB, PubChem, and DrugBank, which are essential for evaluating drug properties comprehensively and optimizing drug discovery pipelines [62], [63].

C. SHORTCOMINGS OF PREVIOUS STUDIES

Past research on computational drug discovery has revealed challenges and limitations that impede their broader applicability and effectiveness across different fields. A significant issue relates to data quality and management; healthcare and biomedical data frequently exhibit high noise levels, inconsistency, and heterogeneity [53], [64]. These affect the reliability of ML models and result in poor performance, especially when datasets are small, imbalanced, or lack standardization. In many cases, the datasets used for drug discovery models are limited, leading to overfitting and difficulty in generalizing new, unseen data [22], [65]. Furthermore, challenges arise in integrating heterogeneous data from various sources, creating bottlenecks in effectively leveraging this information for predictive modeling [66].

Algorithm performance and optimization present another key challenge. ML models, especially DL, offer advanced

capabilities, but they often struggle with overfitting due to the limited data available for training, particularly in drug-target interaction prediction [19], [67]. Many of these models require extensive fine-tuning of complex hyperparameters, which is time-consuming and computationally expensive. This thereby limits the scalability of models to large datasets, an essential requirement for real-world drug discovery applications [68].

Despite their impressive performance, DL models still face limitations. One limitation of DL, along with future perspectives, is that DL can achieve high accuracy due to advances in feature learning. However, this performance depends on having a large training dataset. When data is limited, DL techniques struggle to generalize effectively and often produce biased estimates of model performance. Traditional shallow ML methods may outperform DL models in these cases, as they are less prone to overfitting and require fewer computational resources [69].

A critical issue in interpretability and explainability is the “black-box” nature of many ML and DL models [70], [71]. These models are opaque in their decision-making processes, making it difficult for researchers, clinicians, and regulators to understand the mechanisms behind their predictions [15]. The lack of transparency is especially problematic in drug discovery, where mechanistic insights are needed for regulatory approval and clinical adoption [72]. Recent efforts in explainable AI (XAI) aim to address this issue, but achieving a balance between model complexity and interpretability remains a significant hurdle [3], [73].

Regulatory compliance and ethical considerations further complicate the adoption of AI in drug discovery. AI models raise questions about accountability, fairness, and privacy, especially when handling sensitive patient data [70]. Ensuring that AI models comply with ethical standards and regulatory guidelines is essential for their wider adoption in clinical settings. However, the challenges of meeting these requirements limit the speed at which AI models can be integrated into real-world drug discovery pipelines [17], [18].

The high computational costs of ML models, especially when processing large-scale datasets, require significant resources such as high-performance GPUs or cloud computing platforms [74]. Smaller research institutions often lack access to these resources, limiting their ability to deploy advanced ML techniques effectively. Additionally, the time complexity of training and optimizing DL models restricts their scalability, making them less feasible for practical drug discovery applications [52], [75]. Moreover, validation and experimental confirmation are critical aspects often overlooked in computational studies. Many models are developed and tested using synthetic or pre-clinical data. Still, the actual test of their predictive power lies in experimental validation, particularly in clinical trials [18]. Discrepancies between computational predictions and real-world outcomes often slow the drug development process, as the lack of experimental corroboration reduces the reliability and applicability of

these models [19], [52]. Thus, bridging the gap between computational predictions and experimental results is essential for advancing AI-driven drug discovery.

This study seeks to prioritize critical challenges in drug discovery through the AHP. Systematic ranking of these issues within a scientific framework enables our research to provide valuable insights and methodologies that can meaningfully enhance AI model development in this field.

D. ANALYTIC HIERARCHY PROCESS (AHP)

AHP is a widely recognized decision-making tool that facilitates multi-criteria analysis by decomposing complex problems into a structured, hierarchical model. Initially developed by Thomas Saaty in the late 1970s [76], AHP provides a systematic approach to decision-making that integrates quantitative and qualitative factors, making it highly adaptable across various fields. Its ability to prioritize and weigh diverse criteria has led to its implementation in the healthcare and environmental management sectors, as well as construction and business, where decisions often require balancing competing objectives with significant real-world implications [77].

AHP has proven particularly effective in healthcare and medical decision-making, helping practitioners and policymakers assess and prioritize treatment options, medical technologies, and patient care strategies. Research by Liberatore and Nydick illustrates AHP's capacity to support systematic evaluations based on key criteria such as cost, treatment efficacy, and patient outcomes, thereby enhancing the transparency and consistency of decisions [78]. Similarly, AHP has been instrumental in environmental management, prioritizing sustainable solutions and providing a framework for balancing economic, social, and ecological criteria. Studies underscore AHP's value in facilitating decisions that align with sustainable development goals by enabling stakeholders to analyze and weigh factors critical to environmental preservation [79].

AHP is also extensively applied in construction and engineering [80], [81], where it aids in selecting materials, risk assessment, and resource management. By enabling structured prioritization of factors such as cost, quality, and sustainability, AHP contributes to more robust project planning and execution. Furthermore, in the business sector, AHP supports managerial decision-making by offering a structured approach to evaluate both tangible and intangible factors, such as financial metrics, customer satisfaction, and corporate social responsibility [77]. This structured approach highlights AHP's unique ability to assess criteria that impact strategic business outcomes objectively, thus fostering a holistic approach to decision-making in corporate environments [82].

Beyond traditional sectors, AHP has found critical applications in natural resource management, where it balances the need for resource utilization with conservation imperatives.

Research in this area shows how AHP facilitates the integration of expert opinions, aiding in the development of policies that sustain biodiversity and resource availability [83]. This flexibility has also led to the application of AHP in emerging fields such as technology and innovation management, where structured decision-making is essential to navigate the complexities of modern advancements.

Given the versatility and structured approach of AHP, this study is pioneering its application in AI-driven drug discovery. As AI transforms drug development, AHP systematically prioritizes critical factors like data quality, model interpretability, and ethical considerations. This is the first application of AHP in this field, driven by closely related experts, and it establishes a framework that could significantly improve decision-making in AI-driven pharmaceutical research. The structured and quantifiable insights provided by AHP could lead to more efficient, transparent, and ethical development processes, highlighting its potential as a transformative tool for drug discovery in the era of AI.

III. THE METHODOLOGY OF STUDY

AHP, developed by Thomas Saaty in the 1970s, is a robust multi-criteria decision-making method [76], [84], [85]. This methodology enables decision-makers to structure and prioritize complex issues by organizing them into a hierarchy, allowing for systematic comparisons and quantification of subjective assessments. This study applies AHP to prioritize critical success factors affecting AI-driven drug discovery, employing geometric mean calculations to combine individual judgments into a group judgment matrix. This approach effectively synthesizes inputs from multiple stakeholders, offering a balanced perspective on the importance of each factor [86], [87].

This study uses an AHP hierarchy with three levels. The top level identifies the critical success factors influencing AI-driven drug discovery. The second level presents broad criteria for these factors, while the third level details specific sub-criteria within each category. Unlike traditional AHP studies, this model emphasizes identifying the most influential factors rather than including alternatives.

AHP's capability to address complex decision-making problems makes it well-suited for this study. Its hierarchical structure facilitates a logical breakdown of the problem, and pairwise comparisons help decision-makers evaluate the relative importance of each factor. Using geometric means to aggregate individual assessments ensures that the group judgment matrix reflects a balanced consensus. Although other MCDM techniques like ANP, TOPSIS, or PROMETHEE may be suitable in different contexts [87], [88], [89], AHP's structured framework and proven success in hierarchical analysis make it particularly fitting for this research.

Figure 3 illustrates a flowchart that outlines the research workflow, from defining the research goal to obtaining the results. This visual representation highlights the interconnected steps and the study's systematic approach.

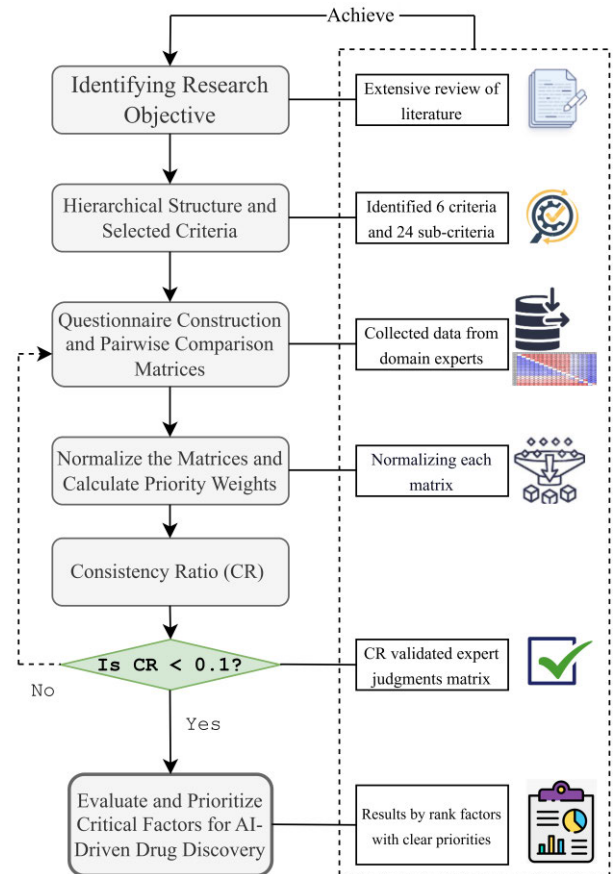


FIGURE 3. Research methodology flowchart.

The research methodology begins by defining the primary goal and developing a hierarchical structure of relevant criteria and sub-criteria. Questionnaires and pairwise comparison matrices are created to capture expert input and evaluate the relationships among the criteria and their sub-criteria. These pairwise comparison matrices are then normalized to calculate priority weights. Afterwards, a consistency ratio (CR) is verified to ensure the comparisons are logically consistent. Finally, the priority scores are aggregated to prioritize the critical success factors. This structured approach enhances the reliability and validity of the study's outcomes.

A. DEFINE THE GOAL

The first and most crucial step in AHP is to define the primary objective. After an extensive literature review, this study establishes to assess and rank the essential factors of success influencing AI-driven drug discovery. Although many studies recognize these factors as challenges, limitations, and concerns in AI-driven drug discovery, none have explored them comprehensively or established a framework for systematically identifying and prioritizing their significance.

This study addresses this gap by introducing a solid AHP-based framework driven by insights from 21 domain experts with extensive knowledge in AI-driven drug discovery. These experts ensure that all aspects and concerns in the domain

TABLE 2. Decision criteria and their sub-criteria for evaluating AI-driven drug discovery.

Criteria	Description	Sub-criterion
DQM	Ensures datasets are accurate, complete, and well-integrated, forming a solid foundation for AI-driven drug discovery. Addresses class imbalances to improve model robustness and predictive accuracy, establishing a reliable basis for successful outcomes.	ACC, CMP, CIH, IHD
APO	Refines AI models to maximize performance across diverse datasets. It emphasizes enhancing generalizability, fine-tuning hyperparameters, addressing overfitting, and leveraging techniques like transfer learning to improve adaptability and efficiency in identifying potential drug candidates.	GEN, HT, CO, TL
IE	AI models must be transparent and their predictions easily understandable, adhering to established scientific principles. Utilizing explainable techniques fosters trust and aids the integration of AI systems into regulatory and practical frameworks, bridging complexity and usability.	XAI, PST, ADK, BBMC
RCEC	It emphasizes the essential need for ethical and regulatory adherence in AI-driven drug discovery. It aims to address issues related to bias, fairness, data privacy, and compliance with standards, ensuring the responsible and equitable implementation of AI.	BF, DPS, EAU, ATS
CES	Optimizes AI systems to handle large-scale datasets and complex tasks effectively. Highlights resource optimization, readiness of computational infrastructure, and exploration of advanced technologies like quantum computing to meet growing demands.	SOS, OOR, IR, FT
VEC	It emphasizes the importance of validating AI predictions through experimental testing and benchmarking. It includes methods for uncertainty estimation and continuous improvement via feedback loops, ensuring the practical applicability and reliability of AI-driven insights.	EV, UE, BM, FL

are captured comprehensively, enabling a systematic evaluation and prioritization process. By defining this goal, the framework integrates expert perspectives to address key challenges, delivering actionable insights that advance the field of AI-driven drug discovery.

B. HIERARCHICAL STRUCTURE AND SELECTED CRITERIA

The study systematically organizes and selects criteria and sub-criteria based on extensive literature reviews addressing key challenges in the domain. Six main criteria and 24 sub-criteria were identified, representing critical success factors. These criteria ensure a comprehensive evaluation framework, balancing scientific insights and practical challenges.

Table 2 outlines the six primary criteria and their corresponding descriptions and sub-criteria (factors), forming a comprehensive framework for evaluating critical success factors in AI-driven drug discovery. These criteria encompass Data Quality and Management (DQM), Algorithm Performance and Optimization (APO), Interpretability and Explainability (IE), Regulatory Compliance and Ethical Considerations (RCEC), Computational Efficiency and Scalability (CES), and Validation and Experimental Confirmation (VEC). Each criterion is further elaborated with specific sub-criteria, detailed in Table 3, including in-depth descriptions and relevant references for each factor.

DQM’s sub-criteria include Accuracy (ACC), Completeness (CMP), Integrity of Heterogeneous Data (IHD), and Class Imbalance Handling (CIH).

The APO criterion addresses key aspects such as Generalization (GEN), Hyperparameter Tuning (HT), Overfitting Control (CO), and Training Learning (TL).

Similarly, the IE criterion comprises Explainable AI (XAI), Black-box Model Challenges (BBMC), Alignment with Domain Knowledge (ADK), and Predictive Sensitivity Testing (PST).

Within the RCEC criterion, the sub-criteria include Adherence to Standards (ATS), Bias and Fairness (BF), Data Privacy and Security (DPS), and Ethical AI Usage (EAU).

The CES emphasizes Infrastructure Readiness (IR), Scalability of Solutions (SOS), Optimization of Resources (OOR), and Future Technologies (FT).

Lastly, the VEC criterion comprises Experimental Validation (EV), Uncertainty Estimation (UE), Benchmark Models (BM), and Feedback Loop (FL).

These factors were identified through a comprehensive literature review based on their significant relevance to challenges, limitations, and unresolved issues highlighted in previous studies on AI-driven drug discovery. While some studies directly address AI methodologies applied to drug development, others provide insights into related processes that influence discovery outcomes. Tables 2 and 3 offer a structured framework, emphasizing the critical role of these criteria in the current research.

Figure 4 illustrates the hierarchical structure of factors influencing AI-driven drug discovery using the AHP method. After establishing criteria and sub-criteria, we developed a structured decision framework to rank and prioritize these factors across three levels systematically:

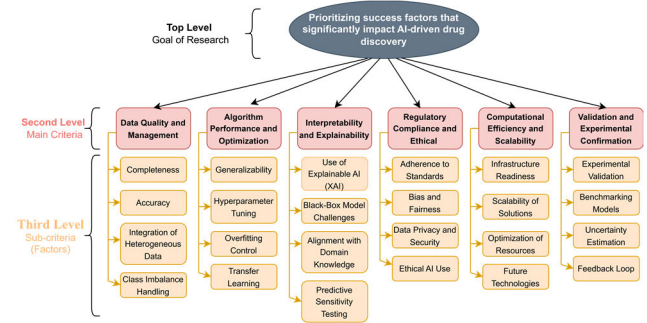


FIGURE 4. Hierarchical levels of AHP for prioritizing critical factors in AI-Driven Drug Discovery.

The top level represents the main objective: prioritizing critical success factors vital for AI-driven drug discovery. The second level organizes six broad criteria: DQM, APO,

TABLE 3. Sub-criteria descriptions in this research.

Sub-criterion	Description	Ref.
CMP	Ensures the dataset is comprehensive, containing all necessary data from various sources, such as omics, clinical trials, and public databases, to enhance model performance.	[12], [90], [7], [91], [92], [9], [6], [93], [4], [94], [95], [96], [97], [98], [71], [99], [98].
ACC	It aims to reduce noise and inaccuracies in data to avoid negative impacts on AI model outputs.	[12], [7], [91], [40], [9], [6], [93], [94], [97]
IHD	combining diverse datasets (e.g., PubChem, ChEMBL) mitigates heterogeneity and ensures consistent data usage.	[7], [92], [40], [6], [38], [93], [100], [101], [4], [97].
CIH	This corrects skewed data distributions, like drug-target interaction data, to prevent biased model predictions.	[90], [7], [12], [102], [103], [14], [97], [104], [71], [105], [99], [98].
GEN	Referring to the model's ability to work effectively across different disease domains and unseen datasets.	[12], [7], [91], [60], [92], [106], [9], [107], [108], [94], [100], [95], [95], [109], [97], [99].
HT	It employs Hyperparameter Tuning methods such as Bayesian optimization to refine model parameters and enhance performance.	[12], [4], [110], [111], [109], [102], [103], [112], [113], [96], [90].
CO	Overfitting is a common issue in AI models and can be reduced using different methods, such as ensemble learning and regularization.	[40], [13], [94], [114], [95], [101], [115], [12], [4], [111], [109], [103], [96], [14], [90], [22].
TL	Uses pre-trained models to overcome the lack of labeled data, allowing better training and optimization for new tasks.	[116], [117], [118], [22], [12], [96], [119], [94], [120], [116].
XAI	Explainable AI improves transparency in AI models through tools like LIME and SHAP that elucidate predictions.	[121], [7], [91], [60], [92], [9], [93], [100], [95], [115], [12], [14], [104].
BBMC	Many ML models are black boxes, posing challenges for explainability in drug discovery.	[7], [91], [60], [9], [6], [100], [95], [115], [12].
ADK	The outputs must consistently align with established scientific knowledge to foster trust and credibility among domain experts.	[122], [123], [124], [125], [126], [127], [128], [12], [95], [129], [130], [101], [115], [4], [127], [131].
PST	Sensitivity analysis ensures the robustness and stability of AI predictions by evaluating their response to variations in input data.	[132], [133], [134], [40], [95], [129], [101], [90], [135].
ATS	AI models must meet pharmaceutical regulatory standards to ensure reliability.	[121], [7], [91], [60], [92], [6], [129], [115], [4], [22].
BF	It addresses algorithmic bias to ensure fairness and prevent discrimination in AI predictions.	[121], [7], [91], [60], [100], [115], [12], [14], [97], [90].
DPS	Ensure compliance with data privacy regulations, safeguarding sensitive patient information in AI-driven drug discovery.	[121], [7], [91], [60], [92], [115], [12], [128], [26].
EAU	Incorporates ethical principles, such as fairness and avoiding discrimination, into AI models to promote responsible use.	[121], [7], [91], [136], [93], [115], [12].
IR	AI applications in drug discovery need to ensure that models access the required computational resources, such as GPUs and cloud platforms, for efficient processing.	[137], [100], [128], [138], [4], [102], [103], [96], [14], [97], [104], [71], [105], [90], [99], [98], [119], [22], [126].
SOS	AI models must scale to multiple datasets and platforms to remain effective while maintaining performance.	[106], [139], [95], [129], [12], [101], [115], [109], [4], [102], [96], [14], [98], [97], [104], [71], [119], [138], [126].
OOR	It involves optimizing computational resources to handle large-scale data processing efficiently.	[129], [101], [111], [102], [103], [96], [14], [104], [71], [90], [98], [138], [126].
FT	Explores cutting-edge technologies to enhance the capabilities of AI models further, focusing on innovative computational advancements and hardware optimizations.	[95], [129], [101], [115], [109], [4], [103], [138].
EV	AI predictions must be rigorously validated using clinical trial data to ensure reliability and enhance trustworthiness in real-world applications.	[7], [91], [92], [40], [140], [93], [94], [100], [129], [101], [115], [12], [14], [104], [90].
BM	Utilizing standard datasets, such as PubChem and ChEMBL, ensures fair evaluation and reliable benchmarking of AI model performance.	[7], [95], [129], [101], [115], [12], [4], [96], [14], [71], [105], [90].
UE	Quantifying uncertainty aids in managing risks and making informed decisions during drug development.	[115], [7], [100], [95], [12], [96], [14], [97], [71], [105], [22].
FL	Implementing feedback loops to update AI models based on experimental outcomes enhances their accuracy and overall performance.	[115], [7], [100], [95], [96], [14], [97], [99], [22], [138].

IE, RCEC, CES, and VEC. The third level details specific sub-criteria within each criterion, facilitating a structured evaluation.

C. QUESTIONNAIRE CONSTRUCT AND PAIRWISE COMPARISON MATRICES

In this research phase, a structured questionnaire was developed to systematically evaluate and prioritize the chosen criteria and sub-criteria, ensuring alignment with the

identified factors. The complete set of questionnaires, organized, captures relevant factors and establishes an evaluation framework. Experienced experts in AI-driven drug discovery were chosen to validate the framework. Their insights ensured the representation of all critical aspects.

Based on expert inputs, pairwise comparison matrices for the criteria and sub-criteria were developed to assess their relative importance. Each criterion was compared with others in pairs, and sub-criteria within each criterion were similarly

evaluated. The decision-makers used Saaty's scale of 1 to 9 to rate the relative importance of elements.

The geometric mean is applied to each pairwise comparison to aggregate these individual judgments into a single group judgment matrix. For example, if multiple stakeholders provide comparisons $c_{ij}^{(k)}$ for criteria i and j , the aggregated group judgment c_{ij} Calculated as:

$$c_{ij} = \left(\prod_{k=1}^K c_{ij}^{(k)} \right)^{\frac{1}{K}} \quad (1)$$

where K is the number of decision-makers, it results in a group judgment matrix C representing a balanced consensus, providing a foundation for further calculations.

The pairwise comparison matrix C for the criteria take the form:

$$C = [c_{ij}] = \begin{bmatrix} 1 & c_{12} & \cdots & c_{1m} \\ c_{21} & 1 & \cdots & c_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & 1 \end{bmatrix}$$

where each c_{ij} reflects the group judgment on the relative importance of criteria i to j , and diagonal values are set to 1, as each criterion is equally essential to itself.

D. NORMALIZE THE MATRICES AND CALCULATE PRIORITY WEIGHTS

The group judgment matrices are normalized to create priority weights, representing the relative importance of each element. For each column in the criteria matrix C , the column sum is computed, and each element is divided by its column sum to form a normalized matrix. The priority weight p_i for each criterion is then calculated as the average of the normalized row values:

$$p_i = \frac{1}{m} \sum_{j=1}^m n_{ij} \quad (2)$$

where n_{ij} denotes the normalized element for c_{ij} . A similar normalization and averaging process is used for each sub-criteria matrix, generating priority weights for factors within each criterion.

E. CONSISTENCY RATIO (CR) VERIFICATION

The consistency ratio (CR) is calculated to confirm the reliability of the group judgments. First, the weighted sum matrix W is derived by multiplying each element in the criteria matrix C by its corresponding priority weight in P . The maximum eigenvalue λ_{max} is obtained by dividing each element in W by its corresponding priority weight p_i and averaging:

$$\lambda_{max} = \frac{1}{m} \sum_{i=1}^m \frac{w_i}{p_i} \quad (3)$$

where w_i represents each element in W . The consistency index (CI) is calculated as:

$$CI = \frac{\lambda_{max} - m}{m - 1} \quad (4)$$

The consistency ratio is derived by dividing CI by a random index (RI):

$$CR = \frac{CI}{RI(n)} \quad (5)$$

A CR of less than 0.1 indicates acceptable consistency, while higher values suggest adjusting the comparisons for more excellent reliability.

F. AGGREGATE PRIORITY SCORES

Once the priority vectors for the criteria P and sub-criteria P_{sk} are established, each success factor's final priority is calculated by multiplying the weight of each sub-criterion by its parent criterion's priority weight. This yields an aggregated score for each factor, reflecting its importance among the 24 factors in achieving the goal.

IV. DISCUSSION AND RESULTS

This study identifies key success factors for AI-driven drug discovery, illustrating their relative importance using a structured AHP framework. This section examines results from pairwise comparisons of criteria and sub-criteria (factors) to optimize AI methodologies in drug discovery.

The study applied the AHP to evaluate six main criteria and their factors to prioritize critical success factors. Table 4 presents the weights and significance of these categories.

Data Quality and Management (DQM) emerged as the most critical, with a weight of 35.7%, and its sub-criterion, Accuracy (ACC), received the highest score at 40.6% of DQM's weight. This highlights the necessity of reliable and accurate data for practical AI applications in drug discovery. Completeness (CMP), Class Imbalance Handling (CIH), and Integration of Heterogeneous Data (IHD) further underline the importance of comprehensive, balanced, and effectively integrated data.

Algorithm Performance and Optimization (APO) ranks second, with Generalizability (GEN) weighted at 52.4%, indicating the need for robust models across diverse datasets. Other factors like Hyperparameter Tuning (HT), Control Overfitting (CO), and Transfer Learning (TL) also stress adaptability, robustness, and efficiency.

Validation and Experimental Confirmation (VEC), at 14.7%, highlights the importance of Experimental Validation (EV) (57.9% of VEC), reinforcing the need for rigorous experimental support for AI predictions. Complementary sub-criteria like Uncertainty Estimation (UE), Benchmarking Models (BM), and Feedback Loop (FL) emphasize confidence quantification and iterative refinement.

Interpretability and Explainability (IE), with a weight of 11.3%, showcases the importance of Explainable AI (XAI) alongside Predictive Sensitivity Testing (PST), Alignment with Domain Knowledge (ADK), and addressing Black-Box Model Challenges (BBMC).

Regulatory Compliance and Ethical Considerations (RCEC) carry a weight of 7.5%. They focus on Bias and Fairness (BF) and assess ethical AI applications along with

TABLE 4. Weights for AI-driven drug discovery Categories and factors.

Criteria	Weight (%)	Sub-criteria	Weight (%)
DQM	35.7%	ACC	40.6%
		CMP	22.9%
		CIH	22.5%
		IHD	14.1%
APO	24.9%	GEN	52.4%
		HT	21.1%
		CO	17.6%
		TL	8.9%
VEC	14.7%	EV	57.9%
		UE	15.0%
		BM	14.3%
		FL	12.8%
IE	11.3%	XAI	53.8%
		PST	16.6%
		ADK	15.7%
		BBMC	14.0%
RCEC	7.5%	BF	50.3%
		DPS	17.9%
		EAU	16.9%
		ATS	15.0%
CES	5.9%	SOS	38.4%
		OOR	25.0%
		IR	21.4%
		FT	15.3%

TABLE 5. Pairwise matrix for Criteria.

Criterion	DQM	APO	IE	RCEC	CES	VEC
DQM	1	1.708	3.694	4.493	5.187	2.150
APO	0.586	1	2.617	3.578	3.948	1.708
IE	0.271	0.382	1	1.898	2.143	0.760
RCEC	0.223	0.280	0.527	1	1.738	0.479
CES	0.193	0.253	0.467	0.575	1	0.478
VEC	0.465	0.586	1.317	2.086	2.091	1

Data Privacy and Security (DPS), Ethical AI Use (EAU), and Adherence to Standards (ATS).

Computational Efficiency and Scalability (CES), the least weighted at 5.9%, stresses the significance of scalable and efficient AI systems, emphasizing Scalability of Solutions (SOS) and Optimization of Resources (OOR), supported by Infrastructure Readiness (IR) and advancements in Future Technologies (FT). These insights present a structured understanding of critical factors and their interdependencies for optimizing AI-driven drug discovery.

The pairwise comparison matrices provide a granular view of the relationships and relative significance among criteria and sub-criteria. Table 5 shows a pairwise matrix for criteria.

Table 6 presents pairwise comparisons of six sub-criteria within each category, highlighting their relationships and relative significance.

TABLE 6. Pairwise matrix comparisons of six sub-criteria.

Data Quality and Management	CMP	ACC	IHD	CIH
CMP	1	0.605	1.718	0.893
ACC	1.652	1	2.976	1.887
IHD	0.582	0.336	1	0.685
CIH	1.120	0.530	1.460	1
Algorithm Performance and Optimization	GEN	HT	CO	TL
GEN	1	2.831	3.374	4.531
HT	0.353	1	1.099	2.933
CO	0.296	0.910	1	2.083
TL	0.221	0.341	0.480	1
Interpretability and Explainability	XAI	BBMC	ADK	PST
XAI	1	4.057	2.955	3.564
BBMC	0.246	1	0.935	0.859
ADK	0.338	1.069	1	0.845
PST	0.281	1.164	1.184	1
Regulatory Compliance and Ethical Considerations	ATS	BF	DPS	EAU
ATS	1	0.370	0.837	0.707
BF	2.703	1	2.993	3.472
DPS	1.195	0.334	1	1.158
EAU	1.413	0.288	0.864	1
Computational Efficiency and Scalability	IR	SOS	OOR	FT
IR	1	0.563	0.768	1.531
SOS	1.775	1	1.554	2.522
OOR	1.302	0.643	1	1.475
FT	0.653	0.397	0.678	1
Validation and Experimental Confirmation	EV	BM	UE	FL
EV	1	5.127	4.011	3.378
BM	0.195	1	1.026	1.313
UE	0.249	0.974	1	1.370
FL	0.296	0.762	0.730	1

Table 7 below summarizes the consistency analysis results for seven pairwise comparison matrices used in this research. Each matrix was evaluated to determine its maximum eigenvalue. λ_{max} and CR Both are essential for validating the reliability of the AHP pairwise comparison process. The λ_{max} indicates the consistency level of the judgments, while the CR Assesses overall consistency, with values below 0.1 deemed acceptable. These results confirm the reliability of the comparisons, ensuring that the priority weights derived from the matrices are valid and appropriate.

This research's results are based on responses from structured questionnaires directed at experts in AI-driven drug discovery. Table 8 and Figure 4 highlight the prioritization of 24 factors, emphasizing the systematic approach needed for addressing the practical crucial success factors.

TABLE 7. Consistency ratio and maximum eigenvalue λ_{max} for pairwise comparison matrices.

Matrix	λ_{max}	CR
Criteria (categories)	6.056	0.0091
Data quality and management criteria	4.009	0.0034
Algorithm performance and optimization criteria	4.040	0.0147
Interpretability and explainability criteria	4.012	0.0046
Regulatory compliance and ethical considerations criteria	4.034	0.0125
Computational efficiency and scalability criteria	4.008	0.0028
Validation and experimental confirmation criteria	4.050	0.0185

TABLE 8. Prioritization of factors in AI-driven drug discovery.

Prioritize	Criterion	Criterion Weight (%)
1	ACC	14.5%
2	GEN	13.0%
3	EV	8.5%
4	CMP	8.2%
5	CIH	8.0%
6	XAI	6.1%
7	HT	5.3%
8	IHD	5.0%
9	CO	4.4%
10	BF	3.8%
11	SOS	2.3%
12	TL	2.2%
13	UE	2.2%
14	BM	2.1%
15	FL	1.9%
16	PST	1.9%
17	ADK	1.8%
18	BBMC	1.6%
19	OOR	1.5%
20	DPS	1.3%
21	EAU	1.3%
22	IR	1.3%
23	ATS	1.1%
24	FT	0.9%

Figure 5 complements the insights from Table 8 by presenting a clear, ranked overview of the crucial factors influencing AI-driven drug discovery. This visualization offers an intuitive comparison that emphasizes key priorities and areas for further improvement, reinforcing the strategic emphasis on AI in pharmaceutical research.

Although experts assigned lower importance to certain factors, their significance should not be overlooked, as many studies highlight their relevance. Factors like ethical AI use, infrastructure readiness, adherence to standards, and future technologies may not directly impact the immediate success of AI-driven drug discovery. Still, they are vital for long-term sustainability and regulatory compliance. Additionally, optimizing resources and ensuring data privacy is essential for maintaining AI systems' efficiency, security, and trustworthiness over time.

The expert assessment reveals that these lower-weight factors, though not primary concerns, function as supportive factors rather than key determinants of AI success in drug discovery. This underscores the need for a prioritization framework that allows decision-makers to allocate efforts and resources efficiently toward the most impactful factors while remaining mindful of secondary, yet potentially influential, considerations.

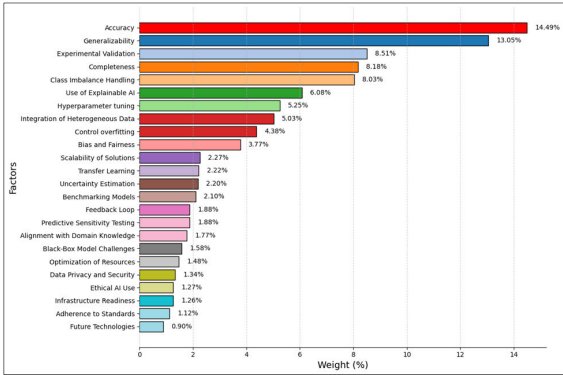


FIGURE 5. Visualization of factors weights in AI-driven drug discovery.

The top 10 factors have been analyzed for their significance and impact. The following section examines these factors, discussing their importance, challenges, and potential effects on optimizing AI-driven drug discovery systems.

A. ACCURACY (ACC) – 14.5%

Accuracy is paramount in AI-driven drug discovery. High accuracy ensures reliable predictions and reduces false positives and negatives during drug candidate screening, vital for cost-effectiveness and clinical success. Inaccurate data and model outputs can impede downstream processes like experimental validation and regulatory approval. Therefore, enhancing data preprocessing, addressing noise, and refining model architecture is crucial for achieving this goal.

B. GENERALIZABILITY (GEN) – 13.0%

Generalizability measures an algorithm's effectiveness across diverse datasets, which is crucial in drug discovery because of the variability in biological data and the complexity of molecular structures.

A generalizable model lessens reliance on large, curated datasets and ensures strong performance in varied contexts, such as rare disease treatments or cross-species analysis. This highlights the importance of thorough training and validation across heterogeneous datasets.

C. EXPERIMENTAL VALIDATION (EV) – 8.5%

Experimental validation provides the bridge between computational predictions and real-world applicability.

This factor's high ranking highlights the necessity of experimental testing to confirm the reliability of AI-generated insights, such as lead compound identification or toxicity

predictions. Strong experimental validation ensures that computational outcomes are theoretical and actionable, reducing the risks associated with clinical trials.

D. COMPLETENESS (CMP) – 8.2%

Completeness refers to the availability and comprehensiveness of datasets used in AI training. In drug discovery, data gaps, whether due to missing molecular properties or incomplete pharmacokinetic profiles, can lead to erroneous predictions. The ranking of this factor signifies the importance of addressing missing data through imputation techniques or by integrating diverse data sources.

E. CLASS IMBALANCE HANDLING (CIH) – 8.0%

Class imbalance handling is critical in drug discovery, where datasets often have a disproportionate representation of active versus inactive compounds. Models trained on imbalanced datasets tend to favor the dominant class, reducing the chances of identifying novel drug candidates. Proper techniques such as oversampling, undersampling, or advanced algorithms like SMOTE (Synthetic Minority Oversampling Technique) are essential for ensuring balanced and unbiased predictions.

F. USE OF EXPLAINABLE AI (XAI) – 6.1%

Explainability in AI models is crucial for trust and transparency, especially in sensitive domains like drug discovery. By providing insights into how predictions are made, XAI enables researchers to understand the underlying decision-making processes, facilitating regulatory approvals and stakeholder acceptance. For example, feature importance analysis can reveal which molecular properties drive efficacy predictions, making XAI an indispensable tool.

G. HYPERPARAMETER TUNING (HT) – 5.3%

Hyperparameter tuning is pivotal in optimizing AI models, as it directly impacts performance metrics such as accuracy and generalizability. The ranking of this factor reflects the need to fine-tune parameters such as learning rates, dropout rates, and the number of hidden layers to achieve optimal results. Automated approaches like grid search or Bayesian optimization can streamline this process, saving time and computational resources.

H. INTEGRATION OF HETEROGENEOUS DATA (IHD) – 5.0%

Drug discovery often combines data from various sources, including chemical properties, biological assays, and clinical trials. Integrating heterogeneous data ensures a holistic view, enabling models to capture complex relationships across domains. This factor's ranking signifies the growing importance of multi-modal data processing in AI-driven drug discovery pipelines.

I. CONTROL OVERFITTING (CO) – 4.4%

Overfitting occurs when an AI model performs well on training data but fails to generalize to unseen data. This issue is particularly relevant in drug discovery, where datasets are often limited in size. Controlling overfitting through regularization techniques, dropout, or cross-validation is critical for building robust and reliable models that can adapt to new challenges.

J. BIAS AND FAIRNESS (BF) – 3.8%

Bias and fairness in AI systems are emerging concerns in drug discovery. Biased models can lead to inequitable outcomes, such as favoring specific drug candidates over others due to dataset imbalances or algorithmic predispositions. This factor emphasizes the importance of addressing biases in training data and ensuring equitable performance across diverse populations, ultimately leading to fairer and more inclusive drug development processes.

The top-ranked factors provide a structured roadmap for advancing AI-driven drug discovery. Enhancing data accuracy, algorithm generalizability, and experimental validation is critical for ensuring reliable and actionable AI outputs. Addressing class imbalance, overfitting, and bias strengthens model fairness and robustness.

This study advances the field by providing a structured, expert-driven prioritization of AI success factors, addressing critical gaps in data integrity, algorithmic transparency, and scalability. Unlike conventional AI adoption strategies focusing on isolated improvements, our approach systematically ranks the most influential criteria, ensuring that AI methodologies are both scientifically valid and practically applicable. By integrating domain expertise into AI evaluation, this research offers a decision-making framework that enhances strategic planning for AI-driven drug discovery, bridging the gap between computational innovations and real-world pharmaceutical requirements.

This study's framework is a decision-making tool, enabling efficient resource allocation and a strategic focus on impactful areas. It also identifies future research opportunities, including ethical considerations, computational efficiency, and leveraging emerging technologies to develop scalable, robust, and ethically sound AI systems for the pharmaceutical domain.

V. IMPLICATIONS, LIMITATIONS, AND FUTURE RESEARCH DIRECTIONS

This study outlines a comprehensive framework for assessing key success factors in AI-driven drug discovery, acting as a strategic resource to enhance decision-making processes in pharmaceutical research. By emphasizing six core criteria, the findings highlight the importance of accurate data, generalizable models, and thorough experimental validation in bridging the gap between computational predictions and their practical applications.

A notable contribution of this research is its systematic prioritization of factors, which offers a clear path for enhancing AI methodologies in alignment with pharmaceutical objectives. Strengthening data management practices, improving algorithm performance, and fostering transparency through interpretability and compliance can address common obstacles, such as data heterogeneity, scalability issues, and trust in AI-based systems. This structured approach equips stakeholders with actionable insights for resource allocation and pipeline optimization.

However, some limitations exist that merit further consideration. The proposed framework was built using a specific set of datasets and expert evaluations, which may limit its adaptability to other contexts within the pharmaceutical domain. AHP is a valuable tool for multi-criteria analysis; however, it assumes that expert judgments are consistent. This assumption may not adequately reflect the complexities of rapidly advancing AI technologies. Furthermore, the dependence on curated datasets emphasizes the importance of considering dynamic, real-world situations and emerging data sources.

Future studies could broaden the framework's scope by incorporating diversity and expanding its application to various stages of drug discovery. Exploring adaptive weighting methods could improve its responsiveness to shifting priorities in AI research and drug development. Integrating domain-specific constraints and advancements, such as generative AI and quantum computing, could further refine the framework's capabilities. Real-world validation across AI-driven drug discovery research as a first step to global initiatives will also be crucial to assess the framework's scalability and robustness. Additionally, introducing multi-objective optimization and adaptive feedback loops could enhance its ability to address complex challenges in real-time settings.

VI. CONCLUSION

This research introduces a framework for identifying key success factors in AI-driven drug discovery and proposes strategies to enhance AI's impact on the pharmaceutical industry. The study highlights critical areas such as data quality, algorithm performance, and experimental validation, underscoring their importance for the effective integration of AI in drug development.

The results suggest that AI's success in drug discovery depends on a balance of technical innovation, ethical considerations, and adaptability. The framework addresses key challenges, including data gaps, algorithmic bias, and regulatory constraints, to meet the growing demand for more efficient and fair therapeutic solutions.

This work contributes to better strategic planning for AI applications in drug discovery by providing actionable insights and a structured approach. It sets the stage for further research to refine AI methods in different pharmaceutical contexts, potentially accelerating drug development and improving patient outcomes. The framework offers a

fresh perspective, informed by expert contributions, and encourages continued innovation, collaboration, and ethical progress.

ACKNOWLEDGMENT

The authors sincerely thank the Deanship of Scientific Research at Imam Mohammad Ibn Saud Islamic University, which was essential for completing this research. This support highlights the University's dedication to advancing scientific innovation and impactful research in emerging fields. They also appreciate the collaborative experts, the conducive environment, and the resources that significantly contributed to this work, including the Study Lounge library in Amman.

REFERENCES

- [1] B. Tanoori and M. Z. Jahromi, "Using drug-drug and protein-protein similarities as feature vector for drug-target binding prediction," *Chemometric Intell. Lab. Syst.*, vol. 217, Oct. 2021, Art. no. 104405.
- [2] A. Khanna, M. El Barachi, S. Jain, M. Kumar, and A. Nayyar, *Artificial Intelligence and Machine Learning in Drug Design and Development*. Hoboken, NJ, USA: Wiley, 2024.
- [3] A. Blanco-González, A. Cabezon, A. Seco-González, D. Conde-Torres, P. Antelo-Riveiro, Á. Piñeiro, and R. Garcia-Fandino, "The role of AI in drug discovery: Challenges, opportunities, and strategies," *Pharmaceutics*, vol. 16, no. 6, p. 891, Jun. 2023.
- [4] J. Vamathevan, D. A. Clark, P. Czodrowski, I. Dunham, E. A. Ferrán, G. Lee, B. Li, A. Madabhushi, P. K. Shah, M. Spitzer, and S. Zhao, "Applications of machine learning in drug discovery and development," *Nat. Rev. Drug Discov.*, vol. 18, no. 6, pp. 463–477, Apr. 2019.
- [5] J. Kim, S. Park, D. Min, and W. Kim, "Comprehensive survey of recent drug discovery using deep learning," *Int. J. Mol. Sci.*, vol. 22, no. 18, p. 9983, Sep. 2021.
- [6] F. C. Udegbe, O. R. Ebulue, C. C. Ebulue, and C. S. Ekesiobi, "Machine learning in drug discovery: A critical review of applications and challenges," *Comput. Sci. IT Res. J.*, vol. 5, no. 4, pp. 892–902, Apr. 2024.
- [7] F. Mirakhor and S. K. Niazi, "Harnessing the AI/ML in drug and biological products discovery and development: The regulatory perspective," *Pharmaceutics*, vol. 18, no. 1, p. 47, Jan. 2025.
- [8] D. B. Catacutan, J. Alexander, A. Arnold, and J. M. Stokes, "Machine learning in preclinical drug discovery," *Nature Chem. Biol.*, vol. 20, no. 8, pp. 960–973, Aug. 2024.
- [9] D. R. Serrano, F. C. Luciano, B. J. Anaya, B. Ongoren, A. Kara, G. Molina, B. I. Ramirez, S. A. Sánchez-Guiraes, J. A. Simon, G. Tomietto, C. Rapti, H. K. Ruiz, S. Rawat, D. Kumar, and A. Lalatsa, "Artificial intelligence (AI) applications in drug discovery and drug delivery: Revolutionizing personalized medicine," *Pharmaceutics*, vol. 16, no. 10, p. 1328, Oct. 2024.
- [10] A. Sahu, J. Mishra, and N. Kushwaha, "Artificial intelligence (AI) in drugs and pharmaceuticals," *Combinat. Chem. High Throughput Screening*, vol. 25, no. 11, pp. 1818–1837, Sep. 2022.
- [11] S. Nag, A. T. K. Baidya, A. Mandal, A. T. Mathew, B. Das, B. Devi, and R. Kumar, "Deep learning tools for advancing drug discovery and development," *3 Biotech*, vol. 12, no. 5, p. 110, May 2022.
- [12] C. Sarkar, B. Das, V. S. Rawat, J. B. Wahlang, A. Nongpiur, I. Tiewsoh, N. M. Lyngdoh, D. Das, M. Bidarolli, and H. T. Sony, "Artificial intelligence and machine learning technology driven modern drug discovery and development," *Int. J. Mol. Sci.*, vol. 24, no. 3, p. 2026, Jan. 2023.
- [13] S. Dara, S. Dhamecherla, S. S. Jadav, C. M. Babu, and M. J. Ahsan, "Machine learning in drug discovery: A review," *Artif. Intell. Rev.*, vol. 55, no. 3, pp. 1947–1999, Aug. 2021.
- [14] G. Obaido, I. D. Mienye, O. F. Egbelowo, I. D. Emmanuel, A. Ogunleye, B. Ogbuokiri, P. Mienye, and K. Aruleba, "Supervised machine learning in drug discovery and development: Algorithms, applications, challenges, and prospects," *Mach. Learn. Appl.*, vol. 17, Sep. 2024, Art. no. 100576.
- [15] Y.-R. Cho and M. Kang, "Interpretable machine learning in bioinformatics," *Methods*, vol. 179, pp. 1–2, Jul. 2020.
- [16] T. Basu, O. Menzer, and S. Engel-Wolf, "The ethics of machine learning in medical sciences: Where do we stand today?" *Indian J. Dermatology*, vol. 65, no. 5, p. 358, 2020.

- [17] H. Shimao, W. Khern-Am-Nuai, K. Kannan, and M. C. Cohen, "Strategic best response fairness in fair machine learning," in *Proc. AAAI/ACM Conf. AI, Ethics, Soc.*, Jul. 2022, p. 664.
- [18] M. Guedj, J. Swindle, A. Hamon, S. Hubert, E. Desvaux, J. Laplume, L. Xuereb, C. Lefebvre, Y. Haudry, C. Gabarroca, A. Aussy, L. Laigle, I. Dupin-Roger, and P. Moingeon, "Industrializing AI-powered drug discovery: Lessons learned from the patrimony computing platform," *Expert Opinion Drug Discovery*, vol. 17, no. 8, pp. 815–824, Aug. 2022.
- [19] L. Cai, J. Chu, J. Xu, Y. Meng, C. Lu, X. Tang, G. Wang, G. Tian, and J. Yang, "Machine learning for drug repositioning: Recent advances and challenges," *Current Res. Chem. Biol.*, vol. 3, Jan. 2023, Art. no. 100042.
- [20] D. Mendez et al., "ChEMBL: Towards direct deposition of bioassay data," *Nucleic Acids Res.*, vol. 47, no. D1, pp. D930–D940, Jan. 2019.
- [21] S. Ali, T. Abuhmed, S. El-Sappagh, K. Muhammad, J. M. Alonso-Moral, R. Confalonieri, R. Guidotti, J. Del Ser, N. Díaz-Rodríguez, and F. Herrera, "Explainable artificial intelligence (XAI): What we know and what is left to attain trustworthy artificial intelligence," *Inf. Fusion*, vol. 99, Nov. 2023, Art. no. 101805.
- [22] C. Cai, S. Wang, Y. Xu, W. Zhang, K. Tang, Q. Ouyang, L. Lai, and J. Pei, "Transfer learning for drug discovery," *J. Med. Chem.*, vol. 63, no. 16, pp. 8683–8694, Jul. 2020.
- [23] A. Darko, A. P. C. Chan, E. E. Ameyaw, E. K. Owusu, E. Pärn, and D. J. Edwards, "Review of application of analytic hierarchy process (AHP) in construction," *Int. J. Construction Manage.*, vol. 19, no. 5, pp. 436–452, Sep. 2019.
- [24] S. Kwon, H. Bae, J. Jo, and S. Yoon, "Comprehensive ensemble in QSAR prediction for drug discovery," *BMC Bioinf.*, vol. 20, no. 1, pp. 1–12, Dec. 2019.
- [25] C. Chakraborty, M. Bhattacharya, S.-S. Lee, Z.-H. Wen, and Y.-H. Lo, "The changing scenario of drug discovery using AI to deep learning: Recent advancement, success stories, collaborations, and challenges," *Mol. Therapy-Nucleic Acids*, vol. 35, no. 3, Sep. 2024, Art. no. 102295.
- [26] A. Smajić, M. Grandits, and G. F. Ecker, "Privacy-preserving techniques for decentralized and secure machine learning in drug discovery," *Drug Discovery Today*, vol. 28, no. 12, Dec. 2023, Art. no. 103820.
- [27] B. Kwak, J. Park, T. Kang, J. Jo, B. Lee, and S. Yoon, "GeoT: A geometry-aware transformer for reliable molecular property prediction and chemically interpretable representation learning," *ACS Omega*, vol. 8, no. 42, pp. 39759–39769, Oct. 2023.
- [28] F. Noor, M. Junaid, A. H. Almalki, M. Almaghrabi, S. Ghazanfar, and M. T. U. Qamar, "Deep learning pipeline for accelerating virtual screening in drug discovery," *Sci. Rep.*, vol. 14, no. 1, p. 28321, Nov. 2024.
- [29] M. Abu-Zanona, B. M. Elzaghmouri, S. M. AbdelRahman, A. A. F. Osman, A. Al-Khateeb, K. Shahzad, and F. Khatoon, "Using artificial intelligence and machine learning approaches to enhance cancer therapy and drug discovery," *J. Ayub Med. College Abbottabad*, vol. 36, no. 1, pp. 183–189, Feb. 2024.
- [30] W. Alzyadat, M. Muhairat, A. Alhroob, and T. Rawashdeh, "A recruitment big data approach to interplay of the target drugs," *Int. J. Adv. Soft Comput. Appl.*, vol. 14, no. 1, pp. 02–13, Apr. 2022.
- [31] S. Y. Ugurlu, "Machine learning applications in drug discovery," *Tech. Rep.*, 2024.
- [32] R. M. Alzhrani, A. H. Almalki, S. L. Alaqel, and S. Alshehri, "Novel numerical simulation of drug solubility in supercritical CO₂ using machine learning technique: Lenalidomide case study," *Arabian J. Chem.*, vol. 15, no. 11, Nov. 2022, Art. no. 104180.
- [33] R. M. Abdallah, H. E. Hasan, and A. Hammad, "Predictive modeling of skin permeability for molecules: Investigating FDA-approved drug permeability with various AI algorithms," *PLOS Digit. Health*, vol. 3, no. 4, Apr. 2024, Art. no. e0000483.
- [34] A. Ahmadi, "Machine learning for drug discovery: Bridging computational science and medicine," *Int. J. BioLife Sci.*, vol. 2, no. 2, pp. 211–220, 2023.
- [35] R. Hajjo, V. Setola, B. L. Roth, and A. Tropsha, "Chemocentric informatics approach to drug discovery: Identification and experimental validation of selective estrogen receptor modulators as ligands of 5-Hydroxytryptamine-6 receptors and as potential cognition enhancers," *J. Medicinal Chem.*, vol. 55, no. 12, pp. 5704–5719, Jun. 2012.
- [36] M. M. Al-Sanea, G. Chilingaryan, N. Abelyan, G. Arakelov, H. Sahakyan, V. G. Arakelov, K. Nazaryan, S. Hussein, G. M. Alazmi, H. E. Alsharari, W. M. Al-Faraj, F. S. Alruwaili, N. Q. Albilasi, T. S. Alsharari, A. A. S. Alsaleh, T. M. Alazmi, A. H. Almalki, N. H. Alotaibi, and M. A. Abdelgawad, "Identification of non-classical hCA XII inhibitors using combination of computational approaches for drug design and discovery," *Sci. Rep.*, vol. 11, no. 1, p. 15516, Jul. 2021.
- [37] S. Al-Showarah, A. Al-Taie, H. E. Salman, W. Alzyadat, and M. Alkhalaileh, "Predicting quality medical drug data towards meaningful data using machine learning," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 8, pp. 1–8, 2023.
- [38] J. K. Yella and A. G. Jegga, "MGATRx: Discovering drug repositioning candidates using multi-view graph attention," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 19, no. 5, pp. 2596–2604, Sep. 2022.
- [39] S. Sunogrot, R. Hamed, H. Abdel-Halim, and O. Tarawneh, "Synergistic interplay of medicinal chemistry and formulation strategies in nanotechnology—from drug discovery to nanocarrier design and development," *Current Topics Medicinal Chem.*, vol. 17, no. 13, pp. 1451–1468, Apr. 2017.
- [40] M. C. Tanay, R. Rathnam, P. V. Krishna, M. Devnath, and N. S. Punna, "Enhanced QSAR modeling for drug discovery: Leveraging advanced computational tools and techniques," in *Proc. IEEE Region Symp. (TEN-SYMP)*, Sep. 2024, pp. 1–6.
- [41] R. Jana, H. M. Begam, and E. Dinda, "The emergence of the C–H functionalization strategy in medicinal chemistry and drug discovery," *Chem. Commun.*, vol. 57, no. 83, pp. 10842–10866, 2021.
- [42] E. H. Weissler et al., "The role of machine learning in clinical research: Transforming the future of evidence generation," *Trials*, vol. 22, no. 1, pp. 1–15, Dec. 2021.
- [43] M. Shah, M. Patel, M. Shah, M. Patel, and M. Prajapati, "Computational transformation in drug discovery: A comprehensive study on molecular docking and quantitative structure activity relationship (QSAR)," *Intell. Pharmacy*, vol. 2, no. 5, pp. 589–595, Oct. 2024.
- [44] J. J. Irwin, T. Sterling, M. M. Mysinger, E. S. Bolstad, and R. G. Coleman, "ZINC: A free tool to discover chemistry for biology," *J. Chem. Inf. Model.*, vol. 52, no. 7, pp. 1757–1768, Jul. 2012.
- [45] S. Kim, J. Chen, T. Cheng, A. Gindulyte, J. He, S. He, Q. Li, B. A. Shoemaker, P. A. Thiessen, B. Yu, L. Zaslavsky, J. Zhang, and E. E. Bolton, "PubChem in 2021: New data content and improved Web interfaces," *Nucleic Acids Res.*, vol. 49, no. D1, pp. D1388–D1395, Jan. 2021.
- [46] A. Al-Hgaish, W. Alzyadat, M. Al-Fayoumi, A. Alhroob, and A. Thunibat, "Preserve quality medical drug data toward meaningful data lake by cluster," *Int. J. Recent Technol. Eng. (IJRTE)*, vol. 8, no. 3, pp. 270–277, Sep. 2019.
- [47] J. Sun, N. Jeliakova, V. Chupakhin, J.-F. Golib-Dzib, O. Engkvist, L. Carlsson, J. Wegner, H. Ceulemans, I. Georgiev, V. Jeliakov, N. Kochev, T. J. Ashby, and H. Chen, "ExCAPE-DB: An integrated large scale dataset facilitating big data analysis in chemogenomics," *J. Cheminformatics*, vol. 9, no. 1, pp. 1–9, Dec. 2017.
- [48] I. D. Mienye and Y. Sun, "Performance analysis of cost-sensitive learning methods with application to imbalanced medical data," *Informat. Med. Unlocked*, vol. 25, Jan. 2021, Art. no. 100690.
- [49] M. Zakariah, S. A. AlQahtani, and M. S. Al-Rakhani, "Machine learning-based adaptive synthetic sampling technique for intrusion detection," *Appl. Sci.*, vol. 13, no. 11, p. 6504, May 2023.
- [50] M. Bhati, T. Virmani, G. Kumar, A. Sharma, and N. Chitranshi, "Deep learning in drug discovery," in *Deep Learning in Personalized Healthcare and Decision Support*. Amsterdam, The Netherlands: Elsevier, 2023, pp. 263–276.
- [51] H. Askar, E. Elgeldawi, H. Aboul Ella, Y. A. M. M. Elshaier, M. M. Gomaa, and A. E. Hassanien, "Deep learning in drug discovery: An integrative review and future challenges," *Artif. Intell. Rev.*, vol. 56, no. 7, pp. 5975–6037, Jul. 2023.
- [52] J.-L. Yu, Q.-Q. Dai, and G.-B. Li, "Deep learning in target prediction and drug repositioning: Recent advances and challenges," *Drug Discovery Today*, vol. 27, no. 7, pp. 1796–1814, Jul. 2022.
- [53] J. Jiménez-Luna, F. Grisoni, and G. Schneider, "Drug discovery with explainable artificial intelligence," *Nature Mach. Intell.*, vol. 2, no. 10, pp. 573–584, Oct. 2020.
- [54] M. Lotfi Shahreza, N. Ghadiri, S. R. Mousavi, J. Varshosaz, and J. R. Green, "A review of network-based approaches to drug repositioning," *Briefings Bioinf.*, vol. 19, no. 5, pp. 878–892, Sep. 2018.
- [55] A. Dalkiran, "Drug-Target interaction prediction by transfer learning for proteins with few bioactive compound data," *Tech. Rep.*, 2024.
- [56] S. Guan and N. Fu, "Class imbalance learning with Bayesian optimization applied in drug discovery," *Sci. Rep.*, vol. 12, no. 1, p. 2069, Feb. 2022.
- [57] N. Rane, S. P. Choudhary, and J. Rane, "Ensemble deep learning and machine learning: Applications, opportunities, challenges, and future directions," *Stud. Med. Health Sci.*, vol. 1, no. 2, pp. 18–41, Jul. 2024.
- [58] M. D. Carmen Legaz-García, J. A. Miñarro-Giménez, M. Menárguez-Tortosa, and J. T. Fernández-Breis, "Generation of open biomedical datasets through ontology-driven transformation and integration processes," *J. Biomed. Semantics*, vol. 7, no. 1, pp. 1–17, Dec. 2016.

- [59] R. Thirumahal and G. Sudhasadasivam, "Semantic integration of heterogeneous healthcare data based on hybrid root linked health record ontology," *Earth Sci. Informat.*, vol. 16, no. 3, pp. 2661–2674, Sep. 2023.
- [60] R. Alizadehsani, S. S. Oyelere, S. Hussain, S. K. Jagatheesaperumal, R. R. Calixto, M. Rahouti, M. Roshanzamir, and V. H. C. De Albuquerque, "Explainable artificial intelligence for drug discovery and development—A comprehensive survey," *IEEE Access*, vol. 12, pp. 35796–35812, 2024.
- [61] W. Shafik, A. F. Hidayatullah, K. Kalinaki, H. Gul, R. Y. Zakari, and A. Tufail, "A systematic literature review on transparency and interpretability of AI models in healthcare: Taxonomies, tools, techniques, datasets, OpenResearch challenges, and future trends," *Tech. Rep.*, 2024.
- [62] A. Uzundurukan, M. Nelson, C. Teske, M. S. Islam, E. Mohamed, J. V. Christy, H. J. Martin, E. Muratov, S. Glover, and D. Fuoco, "Meta-analysis and review of in silico methods in drug discovery-part 1: Technological evolution and trends from big data to chemical space," *Tech. Rep.*, 2024.
- [63] A.-J. Banegas-Luna, J. P. Cerón-Carrasco, and H. Pérez-Sánchez, "A review of ligand-based virtual screening Web tools and screening algorithms in large molecular databases in the age of big data," *Future Medicinal Chem.*, vol. 10, no. 22, pp. 2641–2658, Nov. 2018.
- [64] A. Al-Hgaish, R. Atan, Y. Razali, and M. H. Osman, "Enhancing performance of machine learning models in healthcare: An analytical framework for assessing and improving data quality," in *Proc. Int. Conf. Adv. Inf. Commun. Technol. Comput.*, Oct. 2024, pp. 137–153.
- [65] M. A. Thafar, R. S. Olayan, H. Ashoor, S. Albaradei, V. B. Bajic, X. Gao, T. Gojobori, and M. Essack, "DTiGEMS+: Drug–target interaction prediction using graph embedding, graph mining, and similarity-based techniques," *J. Cheminformatics*, vol. 12, no. 1, pp. 1–17, Dec. 2020.
- [66] J. Wenzel, H. Matter, and F. Schmidt, "Predictive multitask deep neural network models for ADME-tox properties: Learning from large data sets," *J. Chem. Inf. Model.*, vol. 59, no. 3, pp. 1253–1268, Mar. 2019.
- [67] X. Yang, S. Yang, X. Lian, S. Wuchty, and Z. Zhang, "Transfer learning via multi-scale convolutional neural layers for human–virus protein–protein interaction prediction," *Bioinformatics*, vol. 37, no. 24, pp. 4771–4778, Dec. 2021.
- [68] A. Dalkiran, A. Atakan, A. S. Rifaioğlu, M. Martín, R. Çetin-Atalay, A. C. Acar, T. Doğan, and V. Atalay, "Transfer learning for drug-target interaction prediction," *Bioinf.*, vol. 39, no. 1, pp. 103–110, May 2023.
- [69] D. A. Winkler and T. C. Le, "Performance of deep and shallow neural networks, the universal approximation theorem, activity cliffs, and QSAR," *Mol. Informat.*, vol. 36, nos. 1–2, Jan. 2017, Art. no. 1600118.
- [70] P. Pasirja, P. Jha, P. Upadhyaya, M. S. Khan, and M. Chopra, "Machine learning and artificial intelligence: A paradigm shift in big data-driven drug design and discovery," *Current Topics Medicinal Chem.*, vol. 22, no. 20, pp. 1692–1727, Aug. 2022.
- [71] M. K. Tripathi, A. Nath, T. P. Singh, A. S. Ethayathulla, and P. Kaur, "Evolving scenario of big data and artificial intelligence (AI) in drug discovery," *Mol. Diversity*, vol. 25, no. 3, pp. 1439–1460, Aug. 2021.
- [72] P. N. Shiammala, N. K. B. Duraimutharasan, B. Vaseeharan, A. S. Alothaim, E. S. Al-Malki, B. Snekaa, S. Z. Safi, S. K. Singh, D. Velmurugan, and C. Selvaraj, "Exploring the artificial intelligence and machine learning models in the context of drug design difficulties and future potential for the pharmaceutical sectors," *Methods*, vol. 219, pp. 82–94, Nov. 2023.
- [73] D. Minh, H. X. Wang, Y. F. Li, and T. N. Nguyen, "Explainable artificial intelligence: A comprehensive review," *Artif. Intell. Rev.*, vol. 55, no. 5, pp. 3503–3568, Jun. 2022.
- [74] A. Robert, "Machine learning model optimization with GPU acceleration in computational biology," 2024.
- [75] J. Snoek, H. Larochelle, and R. P. Adams, "Practical Bayesian optimization of machine learning algorithms," in *Proc. Adv. Neural Inf. Process. Syst.*, Jan. 2012, pp. 1–46.
- [76] T. L. Saaty, "The analytic hierarchy process (AHP)," *J. Oper. Res. Soc.*, vol. 41, no. 11, pp. 1073–1076, 1980.
- [77] P. Bhambri and S. K. Kautish, "Analytic hierarchy process and business value creation," in *Proc. Using Strategy Anal. Bus. Value Creation Competitive Advantage*. Hershey, PA, USA: IGI Global, Jun. 2024, pp. 54–77.
- [78] M. J. Liberatore and R. L. Nydick, "The analytic hierarchy process in medical and health care decision making: A literature review," *Eur. J. Oper. Res.*, vol. 189, no. 1, pp. 194–207, Aug. 2008.
- [79] P. H. Dos Santos, S. M. Neves, D. O. Sant'Anna, C. H. D. Oliveira, and H. D. Carvalho, "The analytic hierarchy process supporting decision making for sustainable development: An overview of applications," *J. Cleaner Prod.*, vol. 212, pp. 119–138, Mar. 2019.
- [80] H. Hu, S. Jiang, S. S. Goswami, and Y. Zhao, "Fuzzy integrated Delphi-ISM-MICMAC hybrid multi-criteria approach to optimize the artificial intelligence (AI) factors influencing cost management in civil engineering," *Information*, vol. 15, no. 5, p. 280, May 2024.
- [81] I. Masudin, I. Z. Habibah, R. W. Wardana, D. P. Restuputri, and S. S. R. Shariff, "Enhancing supplier selection for sustainable raw materials: A comprehensive analysis using analytical network process (ANP) and TOPSIS methods," *Logistics*, vol. 8, no. 3, p. 74, Jul. 2024.
- [82] J. M. Moreno-Jiménez and L. G. Vargas, "Cognitive multiple criteria decision making and the legacy of the analytic hierarchy process," *Stud. Appl. Econ.*, vol. 36, no. 1, pp. 67–80, May 2019.
- [83] D. L. Schmoldt, *The Analytic Hierarchy Process in Natural Resource and Environmental Decision Making*, vol. 3. Cham, Switzerland: Springer, 2001.
- [84] T. L. Saaty, "Some mathematical concepts of the analytic hierarchy process," *Behaviormetrika*, vol. 18, no. 29, pp. 1–9, Jan. 1991.
- [85] T. L. Saaty and L. G. Vargas, *Models, Methods, Concepts & Applications of the Analytic Hierarchy Process*. Boston, MA, USA: Kluwer, 2001.
- [86] J. E. Leal, "AHP-express: A simplified version of the analytical hierarchy process method," *MethodsX*, vol. 7, Jan. 2020, Art. no. 100748.
- [87] G. Boonthonsatit, S. Vongbunyong, N. Chonsawat, and W. Chanpuypetch, "Development of a hybrid AHP-TOPSIS decision-making framework for technology selection in hospital medication dispensing processes," *IEEE Access*, vol. 12, pp. 2500–2516, 2024.
- [88] S. Bahrami, M. Rastegar, and P. Dehghanian, "An FBWM-TOPSIS approach to identify critical feeders for reliability centered maintenance in power distribution systems," *IEEE Syst. J.*, vol. 15, no. 3, pp. 3893–3901, Sep. 2021.
- [89] X. Niu, Y. Song, and H. Zhu, "Data-driven decision aids for purchasing battery electric vehicles based on PROMETHEE-II methodology," *IEEE Access*, vol. 12, pp. 27931–27946, 2024.
- [90] S. K. Hussin, S. M. Abdelmageid, A. Alkhalil, Y. M. Omar, M. I. Marie, and R. A. Ramadan, "Handling imbalance classification virtual screening big data using machine learning algorithms," *Complexity*, vol. 2021, no. 1, Jan. 2021, Art. no. 6675279.
- [91] P. C. Tiwari, R. Pal, M. J. Chaudhary, and R. Nath, "Artificial intelligence revolutionizing drug development: Exploring opportunities and challenges," *Drug Develop. Res.*, vol. 84, no. 8, pp. 1652–1663, Dec. 2023.
- [92] Y. Qiu, Y. Zhang, Y. Deng, S. Liu, and W. Zhang, "A comprehensive review of computational methods for drug-drug interaction detection," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 19, no. 4, pp. 1968–1985, Jul. 2022.
- [93] C. Selvaraj, I. Chandra, and S. K. Singh, "Artificial intelligence and machine learning approaches for drug design: Challenges and opportunities for the pharmaceutical industries," *Mol. Diversity*, vol. 26, no. 3, pp. 1893–1913, Jun. 2022.
- [94] M. Elbadawi, S. Gaisford, and A. W. Basit, "Advanced machine-learning techniques in drug discovery," *Drug Discovery Today*, vol. 26, no. 3, pp. 769–777, Mar. 2021.
- [95] T. M. Nguyen, T. P. Quinn, T. Nguyen, and T. Tran, "Explaining black box drug target prediction through model agnostic counterfactual samples," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 20, no. 2, pp. 1020–1029, Mar. 2023.
- [96] A. J. Minnich, K. McLoughlin, M. Tse, J. Deng, A. Weber, N. Murad, B. D. Madej, B. Ramsundar, T. Rush, S. Calad-Thomson, J. Brase, and J. E. Allen, "AMPL: A data-driven modeling pipeline for drug discovery," *J. Chem. Inf. Model.*, vol. 60, no. 4, pp. 1955–1968, Apr. 2020.
- [97] E. López-López, E. Fernández-De Gortari, and J. L. Medina-Franco, "Yes SIR! On the structure–inactivity relationships in drug discovery," *Drug Discovery Today*, vol. 27, no. 8, pp. 2353–2362, Aug. 2022.
- [98] C. Wu and L. Chen, "A model with deep analysis on a large drug network for drug classification," *Math. Biosciences Eng.*, vol. 20, no. 1, pp. 383–401, 2022.
- [99] S. Mensa, E. Sahin, F. Tacchino, P. Kl Barkoutsos, and I. Tavernelli, "Quantum machine learning framework for virtual screening in drug discovery: A prospective quantum advantage," *Mach. Learning, Sci. Technol.*, vol. 4, no. 1, Mar. 2023, Art. no. 015023.
- [100] S. Lee, S. Kim, J. Lee, J.-Y. Kim, M.-H. Song, and S. Lee, "Explainable artificial intelligence for patient safety: A review of application in pharmacovigilance," *IEEE Access*, vol. 11, pp. 50830–50840, 2023.
- [101] H. Ahmed, S. Hamad, H. A. Shedeed, and A. S. Hussein, "Enhanced deep learning model for personalized cancer treatment," *IEEE Access*, vol. 10, pp. 106050–106058, 2022.

- [102] A. Mahdaddi, S. Meshoul, and M. Belguidoum, "EA-based hyperparameter optimization of hybrid deep learning models for effective drug-target interactions prediction," *Expert Syst. Appl.*, vol. 185, Dec. 2021, Art. no. 115525.
- [103] S. Bhusare, M. D. Ghorpade, and D. G. Sanap, "Artificial intelligence in drug discovery and development," *Int. J. Res. Publication Rev.*, vol. 6, no. 1, pp. 1096–1106, Jan. 2025.
- [104] Y. Liu, Y. Wang, O. Vu, R. Moretti, B. Bodenheimer, J. Meiler, and T. Derr, "Interpretable chirality-aware graph neural network for quantitative structure activity relationship modeling in drug discovery," in *Proc. AAAI Conf. Artif. Intell.*, Jun. 2023, vol. 37, no. 12, pp. 14356–14364.
- [105] J. Alvarsson, S. Arvidsson McShane, U. Norinder, and O. Spjuth, "Predicting with confidence: Using conformal prediction in drug discovery," *J. Pharmaceutical Sci.*, vol. 110, no. 1, pp. 42–49, Jan. 2021.
- [106] P.-H. Wang, J.-H. Chen, Y.-Y. Yang, C. Lee, and Y. J. Tseng, "Recent advances in quantum computing for drug discovery and development," *IEEE Nanotechnol. Mag.*, vol. 17, no. 2, pp. 26–30, Apr. 2023.
- [107] M.-I. Bittner and S. Farajnia, "AI in drug discovery: Applications, opportunities, and challenges," *Patterns*, vol. 3, no. 6, Jun. 2022, Art. no. 100529.
- [108] P. G. McCabe, S. Ortega-Martorell, and I. Olier, "Benchmarking multi-task learning in predictive models for drug discovery," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2019, pp. 1–7.
- [109] T. R. Noviany, G. M. Idroes, and I. Hardi, "Machine learning approach to predict AXL kinase inhibitor activity for cancer drug discovery using Bayesian optimization-XGBoost," *J. Soft Comput. Data Mining*, vol. 5, no. 1, pp. 46–56, Jun. 2024.
- [110] S. Bonner, I. P. Barrett, C. Ye, R. Swiers, O. Engkvist, C. T. Hoyt, and W. L. Hamilton, "Understanding the performance of knowledge graph embeddings in drug discovery," *Artif. Intell. Life Sci.*, vol. 2, Dec. 2022, Art. no. 100036.
- [111] J. Zhang, Q. Wang, and W. Shen, "Hyper-parameter optimization of multiple machine learning algorithms for molecular property prediction using hyperopt library," *Chin. J. Chem. Eng.*, vol. 52, pp. 115–125, Dec. 2022.
- [112] J. A. Ilemobayo, O. I. Durodola, O. Alade, O. J. Awotunde, A. T. Olanrewaju, O. Falana, A. Ogungbire, A. Osinuga, D. Ogunbiyi, A. Ifeanyi, I. E. Odezuligbo, and O. E. Edu, "Hyperparameter tuning in machine learning: A comprehensive review," *J. Eng. Res. Rep.*, vol. 26, no. 6, pp. 388–395, Jun. 2024.
- [113] L. Colliandre and C. Müller, "Bayesian optimization in drug discovery," in *Proc. High Perform. Comput. Drug Discovery Biomed.* Cham, Switzerland: Springer, Sep. 2023, pp. 101–136.
- [114] A. Lavecchia, "Deep learning in drug discovery: Opportunities, challenges and future prospects," *Drug Discovery Today*, vol. 24, no. 10, pp. 2017–2032, Oct. 2019.
- [115] M. C. Vattikuti, "Improving drug discovery and development using AI: Opportunities and challenges," *Res. J.*, vol. 10, no. 10, pp. 1–24, 2024.
- [116] W. Guo, Y. Dong, and G.-F. Hao, "Transfer learning empowers accurate pharmacokinetics prediction of small samples," *Drug Discovery Today*, vol. 29, no. 4, Apr. 2024, Art. no. 103946.
- [117] R. S. Simoes, V. G. Maltarollo, P. R. Oliveira, and K. M. Honorio, "Transfer and multi-task learning in QSAR modeling: Advances and challenges," *Frontiers Pharmacol.*, vol. 9, p. 74, Feb. 2018.
- [118] Y. Li, Y. Xu, and Y. Yu, "CRNNTL: Convolutional recurrent neural network and transfer learning for QSAR modeling in organic drug and material discovery," *Molecules*, vol. 26, no. 23, p. 7257, Nov. 2021.
- [119] D. Zhuang and A. K. Ibrahim, "Deep learning for drug discovery: A study of identifying high efficacy drug compounds using a cascade transfer learning approach," *Appl. Sci.*, vol. 11, no. 17, p. 7772, Aug. 2021.
- [120] D. Provasi and M. Filizola, "Transfer learning of deep hybrid ligand-based and structure-based drug discovery schemes for improved prediction of opioid ligands' bioactivity," *Biophysical J.*, vol. 123, no. 3, p. 428a, Feb. 2024.
- [121] C. Sissi and M. Palumbo, "Ethical aspects of innovative medical and medicinal sciences," *Medicinal Chem. Res.*, vol. 33, no. 11, pp. 1996–2000, Nov. 2024.
- [122] M. Zinner, F. Dahlhausen, P. Boehme, J. Ehlers, L. Bieske, and L. Fehring, "Toward the institutionalization of quantum computing in pharmaceutical research," *Drug Discovery Today*, vol. 27, no. 2, pp. 378–383, Feb. 2022.
- [123] S. Gao, A. Fang, Y. Huang, V. Giunchiglia, A. Noori, J. R. Schwarz, Y. Ektefaie, J. Kondic, and M. Zitnik, "Empowering biomedical discovery with AI agents," *Cell*, vol. 187, no. 22, pp. 6125–6151, Oct. 2024.
- [124] M. Schaupterl and R. A. Denny, "AI-based protein structure prediction in drug discovery: Impacts and challenges," *J. Chem. Inf. Model.*, vol. 62, no. 13, pp. 3142–3156, Jul. 2022.
- [125] K. Huang, T. Fu, W. Gao, Y. Zhao, Y. Roohani, J. Leskovec, C. W. Coley, C. Xiao, J. Sun, and M. Zitnik, "Therapeutics data commons: Machine learning datasets and tasks for drug discovery and development," 2021, *arXiv:2102.09548*.
- [126] A. Mohan, S. Banerjee, and K. Sekar, "Role of advanced computing in the drug discovery process," in *Proc. Innov. Implementations Comput. Aided Drug Discov. Strateg. Ration. Drug Des.*, 2021, pp. 59–90.
- [127] S. A. Kumar, T. D. A. Kumar, N. M. Beeraka, G. V. Pujar, M. Singh, H. S. N. Akshatha, and M. Bhagyalalitha, "Machine learning and deep learning in data-driven decision making of drug discovery and challenges in high-quality data acquisition in the pharmaceutical industry," *Future Medicinal Chem.*, vol. 14, no. 4, pp. 245–270, Feb. 2022.
- [128] A. Gangwal, A. Ansari, I. Ahmad, A. K. Azad, and W. M. A. Wan Sulaiman, "Current strategies to address data scarcity in artificial intelligence-based drug discovery: A comprehensive review," *Comput. Biol. Med.*, vol. 179, Sep. 2024, Art. no. 108734.
- [129] F. Castiglione, C. Nardini, E. Onofri, M. Pedicini, and P. Tieri, "Explainable drug repurposing approach from biased random walks," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 20, no. 2, pp. 1009–1019, Mar. 2023.
- [130] Y. Ji, L. Zhang, J. Wu, B. Wu, L.-K. Huang, T. Xu, Y. Rong, L. Li, J. Ren, D. Xue, H. Lai, S. Xu, J. Feng, W. Liu, P. Luo, S. Zhou, J. Huang, P. Zhao, and Y. Bian, "DrugOOD: Out-of-distribution (OOD) dataset curator and benchmark for AI-aided drug discovery—A focus on affinity prediction problems with noise annotations," 2022, *arXiv:2201.09637*.
- [131] O. Yildirim, M. Gottwald, P. Schüller, and M. C. Michel, "Opportunities and challenges for drug development: public-private partnerships, adaptive designs and big data," *Frontiers Pharmacol.*, vol. 7, p. 461, Dec. 2016.
- [132] R. Pal, *Predictive Modeling of Drug Sensitivity*. New York, NY, USA: Academic, 2016.
- [133] J. W. Scannell, J. Bosley, J. A. Hickman, G. R. Dawson, H. Truebel, G. S. Ferreira, D. Richards, and J. M. Treherne, "Predictive validity in drug discovery: What it is, why it matters and how to improve it," *Nature Rev. Drug Discovery*, vol. 21, no. 12, pp. 915–931, Dec. 2022.
- [134] E. Ileana Dumbrava, F. Meric-Bernstam, and T. A. Yap, "Challenges with biomarkers in cancer drug discovery and development," *Expert Opinion Drug Discovery*, vol. 13, no. 8, pp. 685–690, Aug. 2018.
- [135] J. Johansson, M. H. Larsson, and J. J. Hornberg, "Predictive in vitro toxicology screening to guide chemical design in drug discovery," *Current Opinion Toxicology*, vol. 15, pp. 99–108, Jun. 2019.
- [136] V. S. Bhargavi, A. Dogra, A. Bhatt, P. Ghildiyal, S. R. Mahendran, and R. Kaur, "Accelerating pharmaceutical research and development through AI-driven drug discovery," in *Proc. 9th Int. Conf. Sci. Technol. Eng. Math. (ICONSTEM)*, Apr. 2024, pp. 1–5.
- [137] O. Spjuth, J. Frid, and A. Hellander, "The machine learning life cycle and the cloud: Implications for drug discovery," *Expert Opinion Drug Discovery*, vol. 16, no. 9, pp. 1071–1079, Sep. 2021.
- [138] S. Gathu, "High-performance computing and big data: Emerging trends in advanced computing systems for data-intensive applications," *J. Adv. Comput. Syst.*, vol. 4, no. 8, pp. 22–35, 2024.
- [139] J. G. Lombardino and J. A. Lowe, "The role of the medicinal chemist in drug discovery—Then and now," *Nature Rev. Drug Discovery*, vol. 3, no. 10, pp. 853–862, Oct. 2004.
- [140] W. Chen, X. Liu, S. Zhang, and S. Chen, "Artificial intelligence for drug discovery: Resources, methods, and applications," *Mol. Therapy-Nucleic Acids*, vol. 31, pp. 691–702, Mar. 2023.



AMIR MOHAMED TALIB received the B.Sc. degree in computer engineering from Technological and Science University, Sudan, in 2006, the M.Sc. degree in computer science from Universiti Putra Malaysia, in 2009, and the Ph.D. degree in software engineering from the Faculty of Computer Science and Information Systems, Universiti Putra Malaysia, in 2012. He is currently an Assistant Professor with the Information Technology Department, College of Computer and Information Sciences, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh, Saudi Arabia. He has more than nine years of teaching experience and about three years of system development experience as a System Developer with Etihad Company, Malaysia. He also teaches system analysis and design and software engineering courses at both undergraduate and graduate levels. His research interests include knowledge management, information and network security, software engineering, and cloud computing. He has also published and written books, articles, and technical papers about his research interests in numerous journals and conference proceedings.



AREEN METIB AL-HGAISH received the B.Sc. degree in software engineering, the B.Sc. degree in communications and computer engineering, and the M.Sc. degree in software engineering. She is currently pursuing the Ph.D. degree in software engineering with the University of Putra Malaysia (UPM).

Her research interests include data quality management, machine learning optimization, and the development of analytical frameworks to enhance artificial intelligence systems' performance. She mainly focused on improving healthcare applications by ensuring the integrity of medical data and exploring innovative AI-driven solutions. She has published research on data quality, machine learning, and AI applications in healthcare, and has actively participated in academic collaborations and conferences. Her contributions reflect a dedication to addressing critical challenges in data-driven technologies and advancing the intersection of data quality and artificial intelligence.



RAZALI YAAKOB (Member, IEEE) received the bachelor's and M.Sc. degrees in computer science from Universiti Putra Malaysia (UPM), in 1996 and 1999, respectively, and the Ph.D. degree from the University of Nottingham, U.K., in 2008.

He is currently an Associate Professor and/or a Lecturer with the Faculty of Computer Science and Information Technology, UPM. He is the Dean of the Faculty. His research interests include artificial neural networks, pattern recognition, and evolutionary gaming computation. He is a member of the Faculty of Intelligent Computing Groups.



RODZIAH BINTI ATAN received the Ph.D. degree in software engineering from the University of Putra Malaysia, in 2005. She is currently with the Faculty of Computer Science and Information Technology, University of Putra Malaysia. She is also an academician responsible for teaching and supervising undergraduate and postgraduate students. Successfully, nine M.Sc. and 11 Ph.D. students graduated as primary supervisors. She specializes in software engineering, process modeling, and cloud computing services. She collaborates with industry and other institutions of higher learning locally and internationally. She has given multiple keynote speeches at international conferences and colloquiums, been appointed as a panel assessor, subject matter expert, local and international journals, and conference reviewer, and was appointed to a position as the Head of the Halal Policy and Management, for seven years, from 2016 to 2023.



ABDULAZIZ ALSHAMMARI received the Ph.D. degree in computer science and informatics from Oakland University, USA. He is currently an Associate Professor with the College of Computer and Information Sciences, Imam Mohammad Ibn Saud Islamic University (IMSIU), Saudi Arabia. His research interests include cybersecurity and machine learning.



ABDULAZIZ ALSAQLI received the Ph.D. degree in information systems from King Saud University, in 2017.

He was the Former Head of the Information Systems Department, College of Computer and Information Sciences, Imam Mohammad Ibn Saud Islamic University. He is currently an Assistant Professor with the College of Computer and Information Sciences, Imam Mohammad Ibn Saud Islamic University. He specializes in software engineering, requirements engineering, and agile software development. His experience includes working as a Consultant with the Ministry of Education, Saudi Arabia, and a member of Saudi Digital Library Development Committee. His current duties include teaching, supervising graduation projects, and supervising master's students. He translated books into Arabic and published them in many ISI-indexed journals and conference papers.



FAHAD OMAR ALOMARY received the B.Sc. degree in electronics engineering from the College of Technology, Riyadh, Saudi Arabia, in 2002, and the M.Sc. degree in computer engineering, the M.Sc. degree in engineering management, and the Ph.D. degree in computer engineering from Florida Institute of Technology, Melbourne, FL, USA, in 2008 and 2013, respectively.

He is currently an Associate Professor with the Information Technology Department, College of Computer and Information Sciences, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh. He has almost ten years of teaching experience at IMSIU and about six years of work experience before teaching in computer engineering and networks. He teaches seminars, computer and ethics, IT fundamentals, and IT professional practices at an undergraduate level. His research interests include knowledge management, information and network security, software engineering, computer engineering, and cloud computing. He has also published and written articles and technical papers about his research interest in numerous journals and conference proceedings.



MOHD HAFEEZ OSMAN received the B.Sc. degree in computer science and the M.Sc. degree in software engineering from Universiti Teknologi Malaysia, Johor Bahru, Malaysia, in 2001 and 2002, respectively, and the Ph.D. degree in computer software engineering from Leiden University, Leiden, The Netherlands, in 2015. His research interests include software engineering and data mining.

...