ITU Journal of Metallurgy and Materials Engineering

ITU/MMEJ

AI Based Advanced Drug Delivery System

Sena D. Aydin[1]

[1]Metallurgical and Materials Engineering Department, Faculty of Chemical and Metallurgical Engineering, Istanbul Technical University, 34469, Maslak, Istanbul, Türkiye.

Abstract: Artificial intelligence (AI) has markedly transformed pharmaceutical research by accelerating the development of target identification, compound screening, molecular optimization, and advanced drug delivery systems.This review addresses the integration of AI and Automated Machine Learning (AutoML) into drug discovery and highlights the emerging role of these technologies in the rational design and optimization of drug delivery materials, from lipid nanoparticles to the molecular structures of materials. AutoML automates fundamental stages of the machine learning process, including data pre-processing, feature engineering, model selection, hyperparameter optimization, and performance evaluation, increasing scalability and reproducibility while reducing reliance on extensive domain expertise. Its ability to process heterogeneous datasets, including omics data, molecular descriptors, and material-related properties, facilitates the optimization of lipid nanoparticles (LNPs), polymeric nanocarriers, and other biomaterials used in advanced drug delivery systems. In addition, advanced models like transfer learning, graph neural networks, transformers, and deep neural network-based virtual screening (DNN-VS) can enhance molecular representation learning to help identify therapeutic agents as well as carriers with useful properties. This review will highlight the increasing convergence of artificial intelligence, bioengineering and materials engineering, demonstrating how data-driven computational strategies are enabling the development of next-generation biomaterials and nanocarrier systems for safer, more efficient, and clinically viable drug delivery applications.

Keywords: AutoML, materials science, artificial intelligence, drug delivery.

Yapay Zeka Destekli Gelişmiş İlaç Dağıtım Sistemleri

Özet: Yapay zeka (AI), hedef tanımlama, bileşik taraması, moleküler optimizasyon ve gelişmiş ilaç dağıtım sistemlerinin geliştirilmesini hızlandırarak ilaç araştırmalarını önemli ölçüde dönüştürmüştür. Bu derleme, yapay zeka ve Otomatik Makine Öğrenimi'nin (AutoML) ilaç keşfine entegrasyonunu ele almakta ve lipit nanopartiküllerden malzemelerin moleküler yapılarına kadar, ilaç dağıtım malzemelerinin rasyonel tasarımı ve optimizasyonunda bu teknolojilerin giderek artan rolünü vurgulamaktadır. AutoML, veri ön işleme, özellik mühendisliği, model seçimi, hiperparametre optimizasyonu ve performans değerlendirmesi dahil olmak üzere makine öğrenimi sürecinin temel aşamalarını otomatikleştirerek ölçeklenebilirliği ve tekrarlanabilirliği artırırken, kapsamlı alan uzmanlığına olan bağımlılığı azaltır. Omik veriler, moleküler tanımlayıcılar ve malzemeyle ilgili özellikler dahil olmak üzere heterojen veri kümelerini işleme yeteneği, gelişmiş ilaç dağıtım sistemlerinde kullanılan lipit nanopartiküllerinin (LNP'ler), polimerik nano taşıyıcıların ve diğer biyomalzemelerin optimizasyonunu kolaylaştırır. Buna ek olarak, transfer öğrenimi, grafik sinir ağları, transformatörler ve derin sinir ağı tabanlı sanal tarama (DNN-VS) gibi gelişmiş modeller, moleküler temsil öğrenimini geliştirerek, terapötik ajanların yanı sıra yararlı özelliklere sahip taşıyıcıların da tanımlanmasına yardımcı olabilir. Bu derleme, yapay zeka, biyomühendislik ve malzeme mühendisliği arasındaki artan yakınlaşmayı vurgulayacak ve veriye dayalı hesaplamalı stratejiler, daha güvenli, daha verimli ve klinik açıdan uygulanabilir ilaç dağıtım uygulamaları için yeni nesil biyomalzemelerin ve nano-taşıyıcı sistemlerin geliştirilmesini mümkün kılmaktadır.

Keywords: AutoML, malzeme bilimi, yapay zeka, ilaç iletimi.

Review Article

Corresponding Author: Sena D.Aydin, E-mail: aydinsen21@itu.edu.tr

Reference: Aydin, S. D. (2026), AI Based Advanced Drug Delivery System, *ITU Journal of Metallurgy and Materials Engineering,3(2) 56-60.DOI: 10.65774/itummej.1967008*

Submission Date: 9 June 2026
Online Acceptance: 22 July 2026
Online Publishing: 29 July 2026

1.Introduction

The evolution of Artificial Intelligence (AI) has led to the automation of complex processes across various technological domains. In the pharmaceutical field, AI has become a transformative technology in drug discovery by accelerating target identification, compound screening, lead optimization, and the design of advanced drug delivery systems such as lipid nanoparticles (LNPs). However, developing effective AI-driven systems remains challenging due to the complexity of biological data and the high computational cost associated with model development. The design and optimization of drug delivery materials require careful consideration of physicochemical properties, including particle size, surface characteristics, encapsulation efficiency, and stability, making materials engineering an integral part of advanced drug delivery system development. Lipid NPs have several applications in the development of vaccines, gene therapy, and protein replacement but may demonstrate toxicity, less efficient delivery, and immune activation [1]. AutoML facilitates automation of certain stages of the AI pipeline:data preprocessing, model selection, hyperparameter tuning, and evaluation. Once the data are curated, days of work are spent on model identification or finalizing an activity prediction, which constitutes an error process and depends on expertise. Machine Learning (ML) techniques are appealing to the pharmaceutical industry due to their automated nature, predictive capabilities, and the consequent expected increase in efficiency [2]. Apart from small molecule discovery, such methods also facilitate the design and optimization of lipid nanoparticles (LNPs) formulations. Conventional methods such as Graph Neural Network (GNN) and transformer models tend to be very computationally intense; thus, a transfer learning approach is used to lay off some of those computational requirements. In such a case, generation models using SMILES have been found to be a potential method to enable AI learn chemical syntax and create structures of molecules for advanced drug delivery systems. Simplified Molecular Input Line Entry System Based (SMILES-based) networks like Recurrent Neural Networks (RNNs), Variational Autoencoder (VAEs), and Generative Adversarial Network (GAN) have openly showcased the ability to learn the grammar of molecules and generate syntactically valid molecules by sampling new molecules with desired properties, these models can be used for de novo drug design. The RNN tries to learn the sequential dependency in molecular structures from SMILES string representations, whereas GANs or VAEs are used to diversify and optimize the chemical space. The limitations being generated in this way have opened up thoughts about systems able to learn molecular representations and generate new molecules having the same drug-like properties. The release of regulatory guidance on the potential use of Bayesian statistics in medical device development and regulatory endorsement of the Bayesian methods for early clinical development has further fueled the enthusiasm for Bayesian applications to drug development issues, including those previously considered intractable [3]. Innovative techniques for drug–target interaction (DTI) prediction are overcoming traditional limitations. To address the challenge of limited training data and poor generalization, ML based on pre-trained protein language models have significantly enhanced extrapolation capabilities by analyzing the structural determinants of molecular recognition [4].

2. Related Work

The integration of AI into pharmaceutical research has been extensively studied, with significant progress in areas such as drug-target interaction, prediction, virtual screening, de novo molecule generation, and Absorption, Distribution, Metabolism, Excretion, and Toxicity (ADMET) property estimation. Early works primarily focused on manually constructed pipelines where domain experts curated features and selected machine learning models for specific tasks. Even though effective, the limitations included scalability and reproducibility over various types of data sets. On the other hand, deep learning algorithms have the advantage of working on big and different data sets, an aspect that is very important in learning nonlinear patterns from biomedical data. The advantage of the manual way is that good model interpretability usually follows, provided one can quantify from the trained model the strength of the influence of each feature (for instance, it is ensured when using linear regression models) [5]. An important note in this process is using descriptive simplified molecular-input line-entry system (SMILES) nomenclature in much of the algorithms regarding de novo drug design and discovery [6].The progress in deep learning has made it possible to design representation learning techniques that can capture subtle biological patterns. Deep Learning has rapidly emerged and has attracted an increasing attention in the pharmaceutical industry and the academic research community, as a promising tool in drug research and development [7]. Graph Neural Network (GNN) has achieved good performance on learning materials from molecular structure, modeling atoms and bonds as nodes and edges. In the same logic, transformer-based architectures Molecule Attention Transformer (MAT) and ChemBERTa have paved the way for context-aware molecular representation models with SMILES or protein sequences representation. These architectures have shown great potential in predicting Lipid Nanoparticle (LNPs) formulations characteristics; e.g. encapsulation efficiencies and particle sizes, using the physical/chemical descriptors as a bridge between Artificial Intelligence and Materials Engineering. However, the majority of these approaches are limited with respect to the workflow is initialized with a data gathering step, which includes the compilation of molecular structures along with their experimental biological of two types of data sets: training data set and test data set. Upon curation, these are subjected to feature extraction using cheminformatics packages such as RDKit or DeepChem, to transform molecular representations (SMILES, InChI) into numerical descriptors or embeddings. Consequently, a deep neural network model was then here termed DNN-VS (Deep Neural Networks for Virtual Screening), was trained on the processed training data.

Figure 1 shows how the VAE–GAN architecture combines the property of representation learning of Variational Autoencoder (VAE) with the property of generating realistic samples of Generative Adversarial Network (GAN). First, the encoder q(y|x) takes molecular descriptors or structural representations of the known compounds as input and encodes them in a low-dimensional latent space. Instead of assigning one latent vector, the encoder learns the distribution, which provides the opportunity to explore the space of chemical meaning of the molecular representations. One of such latent vectors is chosen and decoded into the molecule, retaining its main chemical properties. Additionally, the generator of the GAN architecture takes randomly chosen latent vectors of the prior distribution p(z) to generate new molecules. Generated and reconstructed molecules are then classified by the discriminator, which tries to determine which of the molecules is real and which is fake. Through adversarial training, the generator gradually improves its ability to produce realistic compounds escalating for the discriminator to recognize.

From a materials engineering perspective, such an architecture allows efficient exploration of the design space for functional biomaterials and drug carriers. Due to understanding of complex structure-property relations, the model will allow for fast design and optimization of lipid nanoparticles, polymeric drug carriers, and other advanced

biomaterials with optimized physicochemical properties, enhanced biocompatibility, stability, and drug release profile. Thus, generative models using AI technologies become a promising computational method for rational design of advanced drug carriers. While models such as variational autoencoders(VAEs) and flow-based methods have demonstrated success in generating valid and unique molecules, they often struggle with more complex, drug-like compounds highlighting limitations in their applicability to real-world drug design.In recent years, with advancements in computational power and data availability, large language models (LLMs) have demonstrated exceptional performance in various areas, including content comprehension and reasoning [8].In particular, this has relevance to LNP design, where labelled experimental data sets remain sparse, and transferring from larger molecular libraries presents an important way forward. Merging transfer learning with AutoML now stands as a very exciting pathway to optimize feature extraction and model generalization at reduced computational costs. AI has come to a crossroads and holds the potential of ushering in innovation in widespread ways toward drug discovery and delivery, especially in emergencies such as the COVID-19 pandemic. A broad constellation of AI-based methodologies was pursued to hasten the therapeutics search. Among others, stand GNNs and knowledge graph-based approaches, which show the greatest promise for properly capturing complex biomedical interrelations and hence providing steerage toward drug repurposing.

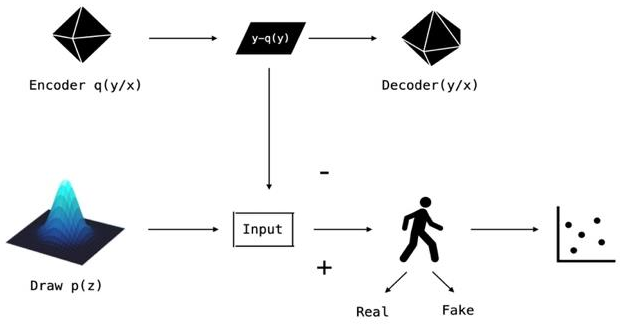

Figure 1. Schematic representation of a hybrid Variational Autoencoder–Generative Adversarial Network (VAE–GAN) architecture for molecular generation in AI-assisted drug discovery. The encoder converts molecular representations into a latent space, while the decoder reconstructs molecular structures from the learned latent variables. In parallel, randomly sampled latent vectors are used to generate new molecular candidates, and the discriminator distinguishes generated molecules from real molecular data. Through adversarial learning, the model improves the validity, diversity, and chemical realism of the generated structures, supporting the computational design and optimization of advanced drug delivery materials. Source: Created by the author.

Şekil 1. Yapay zeka destekli ilaç keşfinde molekül üretimi için kullanılan hibrit Varyasyonel Otokodlayıcı–Üretici Çekişmeli Ağ (VAE–GAN) mimarisinin şematik gösterimi. Kodlayıcı (encoder), moleküler gösterimleri gizil (latent) uzaya dönüştürürken, kod çözücü (decoder) öğrenilen gizil değişkenlerden moleküler yapıları yeniden oluşturur. Paralel olarak, rastgele örneklenen gizil vektörler yeni molekül adaylarının üretilmesinde kullanılır ve ayırt edici ağ (discriminator), üretilen molekülleri gerçek moleküler verilerden ayırt eder. Çekişmeli öğrenme (adversarial learning) sayesinde model, üretilen yapıların geçerliliğini, çeşitliliğini ve kimyasal gerçekçiliğini artırarak gelişmiş ilaç taşıyıcı malzemelerin hesaplamalı tasarımını ve optimizasyonunu destekler. Kaynak: Yazar tarafından oluşturulmuştur.

A medical knowledge graph was constructed by embedding five types of biomedical entities drugs, genes, diseases, channels, and side effects along with nine distinct relationships. GCNs augmented with attention mechanisms were utilized to extract topological features and generate a drug-disease prediction matrix. This strategy enabled the identification of host-targeting drugs capable of disrupting disease-relevant molecular pathways. The method based on the Knowledge Graph Embedding (KGE) maps the entities and relationships in the knowledge graph to a low-dimensional continuous vector space, which can retain the inherent characteristics of the knowledge graph and alleviate the feature sparse problem that may be faced in the application of the knowledge graph. The training process is divided into multiple stages [9]. Unlike Structure-Based Methods, AI based methodologies enable the integration of multiple types of biomedical data to provide a Systems-Level understanding of how diseases work. In addition, AI-driven materials engineering methodologies specifically lipid nanoparticle (LNP) design represent new frontiers with respect to the rational formulation of nucleic acid delivery vehicles, via predictive modelling and high throughput virtual screening.

3. Methods

Arguably, one of the most thrilling applications of Deep Neural Networks in drug discovery is the development of novel chemical entities. The integration of deep learning models, notably Variational Autoencoders (VAEs), has paved the way for the mapping of chemical structures, described as SMILES strings (a molecular structure notation), into a continuous latent space. Subsequent to its training, this model is able to generate new molecules purely by randomly sampling the latent space and decoding the randomly sampled points back into SMILES strings.

This sampling process gives researchers the freedom to design molecules with specific desired properties, such as improved binding affinity or drug- likeness. RNNs, well-suited to sequence systems such as SMILES strings, have also been successfully used to generate valid chemical structures. RNNs are capable of learning the probability distribution of characters in a SMILES string and applying this knowledge to the generation of new, structurally sound molecules. The networks are trained on a large dataset of molecules, and RNNs can autonomously generate new molecules that may possess desirable pharmacological properties. Unfortunately, RNNs suffer from the problem of vanishing gradients, which significantly limits their ability to work with long sequences [10].

3.1. Model Training and Setup

A typical VAE model has two main components encoder and the decoder. The encoder compresses input SMILES strings into some continuous latent space, and the decoder learned to reconstruct them back into SMILES sequences from the standpoint of sampling points from the latent space it generated new molecules that had the properties learned from the original dataset. The same approach is applied when formulating Lipid Nanoparticles (LNPs) using Latent Space Frameworks.The GAN model uses two competing networks, a discriminator and a generator. The discriminator judged whether the new SMILES sequences that the generator tried to create were valid. By using a large number of SMILES molecular structures as training sets, the RNN model they trained can automatically generate molecular structures with high drug-like properties and the efficacy is as high as over 90%, and the diversity of molecular structures obtained is basically the same as that of the training set [11]. The RNN model was designed using stacked Long Short-Term Memory (LSTM) layers that are mostly designed for sequential data like

SMILES. Thus, the Long Short-Term Memory (LSTM) layers in the model preserved long-term dependencies from one atom in a molecule to the next atom. It tokenized the data at the character-level so that the model could learn structural rules directly from raw SMILES syntax. It also used batch normalization and dropout layers in some trials to increase generalizability.

3.2. Training and Sampling Strategy

The training approach of SMILES-based generative methods, mainly based on RNN and GAN representations, will help to optimize both syntactic correctness and chemical meaning. In this work, SMILES representations of molecular structures are initially tokenized and encoded to numerical sequences so that those can be used as input to the generator model. Importantly, RNN-based generators take advantage of the character-level embedding to maintain the sequential properties of the SMILES format.

Models are trained using 'teacher forcing' to predict the next token, under the objective of minimizing categorical cross-entropy loss between predicted and actual token sequences.

To increase the diversity and novelty of molecules produced with the generative models, we utilize a GAN framework that consists of a generator (RNN or Transformer) that is trained to produce valid SMILES strings and a discriminator that can separate real training set and generated molecules. This is an example of how generative artificial intelligence can be applied across computational chemistry and materials science domains to support data driven discovery of delivery systems along with therapeutic molecules they carry

4. Evaluation of Generated Molecules

Assessing the quality, relevance, and usefulness of molecules generated by AI is a critical step in drug discovery. Evaluating these generated molecules requires evaluating more than parsing for syntactic correctness but biological significance, chemical novelty, and drug-likeness. For a more comprehensive evaluation, a multi-metric framework is typically used, focusing on all the following aspects: validity, uniqueness, novelty, diversity, desirability, and synthetic feasibility. Furthermore, domain-relevant evaluations, such as biological activity predictions and synthetic accessibility, must be considered when isolating molecules that have the maximum chance of success in the drug development process. The evaluation criteria are usually categorized into the following dimensions:

1. Validity,ensuring the molecules adhere to chemical rules.
2. Distinctiveness and originality verifying if the model creates unique, never-before-seen structures.
3. Variety assessing how much of the chemical space has been mapped out.
4. Drug-likeness and other major characteristics measuring parameters such as molecular weight, logP, and general appropriateness towards drug development.
5. Suggested biological activity predicting the probability of a molecule binding to certain protein domains.
6. Ease of synthesis evaluating the difficulty of producing the molecules in a laboratory setting.
7. Benchmarking assessing the model's performance against established databases such as GDB-13 or ChEMBL.

These batches are in the form of a matrix with a row for each encoded SMILES string and appended with end tokens as padding. Using smaller training set sizes (10,000 and 1000) further highlights the data augmentation effect of ran-domized SMILES and enables training models that are able to generate 62% of GDB-13 with only a sample comprising 0.001% of the database. When training models on a ChEMBL training set, randomized SMILES models have a much bigger output domain of molecules in the same range of physicochemical properties as the canonical SMILES models [12].

5. Future Challenges

Integrating AI into drug discovery can transform the pharma industry by speeding up the research for new drugs and enhancing the processes involved in developing them. Even with the advancements in AI technologies, there is still more work to be done in order to fully capitalize on AI's capabilities. Often, both the target and the biopharmaceutical are less amenable to structure–activity studies and require significantly different evaluation approaches to those used in the development of small-molecule drugs [13]. Clearly, our increasing ability to develop, deploy, and continuously improve AI methodologies at scale bodes well for the future of AI-driven advanced drug discovery, leading to better drugs to improve people life quality and decreasing cost compare the traditional methods.

5.1. Quality and Integration

AI models require training on high-quality, curated data with sufficient diversity. Within the domain of drug discovery, the biological, chemical, and clinical datasets available are often isolated and incomplete. Gaps in datasets, data collection mistakes, and biases built into the design of experiments all severely limit the ability to train reliable AI models for drug delivery systems. Furthermore, attempting to merge data from genomic, proteomics, and clinical-trial databases is exceedingly difficult due to the myriad of formats, scales, complexities, and even the breadth of disciplines covered by these databases. Meeting this need calls for unwavering patronage of empirical protocols that describe spanning processes at the data level. Moreover, editing algorithms capable of processing imperfect datasets, unbalanced data underused data will strengthen accuracy and reliability.

5.2. Ethical and Regulatory Considerations

Ethics has always been a major issue in any engineering field. Incorporation of Artificial Intelligence (AI) in drug discovery presents a serious ethical and regulatory issue. Use of AI tools in a clinical setting is strictly regulated by law to safeguard the safety of patients. Hence, patient consent must be acquired whenever personal data is being gathered, while at the same time giving utmost respect to patients' rights to information, autonomy, and refusal of participation. Moreover, use of biomedical data must adhere to stringent privacy and security requirements.

Moreover, AI tools could inadvertently perpetuate biases in healthcare, resulting in discrimination in terms of access to care and decision-making processes. Thus, future AI frameworks should be developed such that they are capable of promoting fairness, transparency, reliability, and accountability. It is important for regulatory agencies to also set up guidelines for validating and approving AI-based drug discovery systems to ensure that such systems meet the highest standards of safety and efficacy. Despite the high potential of AI in speeding up drug innovation, there is a need for ethics to catch up with technology. In order to better define the ethical responsibilities of artificial intelligence in drug design, complete laws, regulations and regulatory mechanisms need to be established. These laws and regulations should clarify the status and role of AI in drug design, as well as the responsibilities and obligations of all

relevant parties. At the same time, the regulatory mechanism should conduct comprehensive supervision and evaluation of the decision-making process of AI to ensure its legality and compliance [14].

6. Conclusion

The combination of AI, AutoML, and sophisticated deep learning architectures is radically transforming the process of drug discovery in pharmaceutics. Apart from being pure predictors, these computational techniques are developing into intelligent design systems to enable the acceleration of molecular discovery, optimization of lead compounds, and data-driven decision making during all phases of drug development. The findings compiled in this review prove that the use of molecular representations based on SMILES combined with such architectures as VAEs, GANs, RNNs, transformer models, and GNNs has significantly boosted the exploration of chemical space, increasing the efficiency of molecular generation and virtual screening.

It is noteworthy that the applications of the abovementioned technologies do not confine themselves to traditional small-molecule discovery only. The computational approaches powered by AI are becoming increasingly relevant for the domain of materials engineering, as they provide the means of rational design and optimization of lipid nanoparticles, polymeric drug carriers, and other biomaterials applied in modern drug delivery systems. The determination of the complicated structure–property relations allow enhancing important characteristics of materials, which include particle size, encapsulation efficiency, physicochemical stability, biocompatibility, and drug release. On the other hand, the extensive use of artificial intelligence in pharmaceutical science research continues to face obstacles linked to issues of heterogeneity in data integration, model explainability, data set quality, and compliance with regulations. Such problems suggest that further developments will require not only the improvement of the algorithms but also creating of appropriate data infrastructures, validation procedures, and ethical frameworks for the successful application of artificial intelligence technologies that are capable of generating reliable results. In summary, the information provided in this paper demonstrates that the use of AI and AutoML approaches can be considered as paradigms of new technologies that are redefining not only the discovery of molecules but also the engineering of new generation of therapeutic materials. The integration of such technologies with pharmaceutical sciences and materials engineering is providing a computational framework for the design of safe and personalized therapeutic systems.

7. References

[1] Kapoor, D. U., et al. (2024). AI-driven design and optimization of nanoparticle-based drug delivery systems. *Science, Engineering and Health Studies*, 24010003–24010003.

[2] Kolluri, S., Lin, J., Liu, R., & others. (2022). Machine learning and artificial intelligence in pharmaceutical research and development: A review. *AAPS Journal, 24*, 19. https://doi.org/10.1208/s12248-022-00681-5

[3] Yang, H., & Novick, S. (2019). Bayesian analysis with R for drug development: Concepts, algorithms, and case studies (1st ed.). Chapman and Hall/CRC. https://doi.org/10.1201/9781315100388

[4] Wang Q, Sun B, Yi Y, Velkov T, Shen J, Dai C, Jiang H. Progress of AI-Driven Drug-Target Interaction Prediction and Lead Optimization. Int J Mol Sci. 2025 Oct 15;26(20):10037. doi: 10.3390/ijms262010037. PMID: 41155330; PMCID: PMC12563295.

[5] Réda, C., Kaufmann, E., & Delahaye-Duriez, A. (2020). Machine learning applications in drug development. *Computational and Structural Biotechnology Journal*, 18, 241–252. https://doi.org/10.1016/j.csbj.2019.12.005

[6] Patel, L., Shukla, T., Huang, X., Ussery, D. W., & Wang, S. (2020). Machine learning methods in drug discovery. *Molecules*, 25(22), 5277. https://doi.org/10.3390/molecules25225277

[7] Bahi, M., & Batouche, M. (2018). Deep learning for ligand-based virtual screening in drug discovery. In *2018 3rd International Conference on Pattern Analysis and Intelligent Systems (PAIS)* (pp. 1–5). IEEE. https://doi.org/10.1109/PAIS.2018.8598488

[8] Wang, B., Zhang, T., Liu, Q., Sutcharitchan, C., Zhou, Z., Zhang, D., & Li, S. (2025). *Elucidating the role of artificial intelligence in drug development from the perspective of drug-target interactions*. Journal of Pharmaceutical Analysis, 15, 101144. https://doi.org/10.1016/j.jpha.2024.101144

[9] Che, M., Yao, K., Che, C., Cao, Z., & Kong, F. (2021). Knowledge-Graph-Based Drug Repositioning against COVID-19 by Graph Convolutional Network with Attention Mechanism. *Future Internet*, *13*(1), 13. https://doi.org/10.3390/fi13010013

[10] Grechishnikova D. Transformer neural network for protein-specific de novo drug generation as a machine translation problem. Sci Rep. 2021 Jan 11;11(1):321. doi: 10.1038/s41598-020-79682-4. PMID: 33432013; PMCID: PMC7801439.

[11] Kong W, Hu Y, Zhang J, Tan Q. Application of SMILES-based molecular generative model in new drug design. Front Pharmacol. 2022 Oct 13;13:1046524. doi: 10.3389/fphar.2022.1046524. PMID: 36313306; PMCID: PMC9606214.

[12] Arús-Pous, J., Johansson, S., Prykhodko, O., Bjerrum, E., Tyrchan, C., Reymond, J.-L., Chen, H., & Engkvist, O. (2019). *Randomized SMILES strings improve the quality of molecular generative models*. Journal of Cheminformatics, 11, Article 71. https://doi.org/10.1186/s13321-019-0393-0

[13] Bhogal, N. (2009). 21st century drug development: Advances, opportunities and challenges. *Alternatives to Laboratory Animals,37*(4),335–339. https://doi.org/10.1177/026119290903700401

[14] Chen, T. (2024). Artificial intelligence and drug design: Future prospects and ethical considerations. *Computational Molecular Biology, 14*(1), 9–19. https://doi.org/10.5376/cmb.2024.14.0002