AMERICAN JOURNAL OF MANAGEMENT AND IOT MEDICAL COMPUTING
Peer Reviewed, Referred & Indexed Journal
E-ISSN: 3069-0110
Vol.5, No.2(1) (2026)
www.ajmimc.com

A Trust-Embedded Learning Architecture for Discovering Alternative Drug Indications with Verifiable Computational Integrity

Ch. Jyothi[1*], Koyilakonda Sneha[2], Nenavath Himabindu[2], Midisinametla Aravind[2]

[1]Assistant Professor, [2]UG Student, [1,2]Department of Computer Science and Engineering

[1,2]Kommuri Pratap Reddy Institute of Technology, Ghanpur, Ghatkesar, 501301, Telangana, India.

*Correspondence: Ch. Jyothi (vampu.jyothi9@gmail.com)

ABSTRACT

Drug repurposing has emerged as an effective strategy in modern healthcare, enabling researchers to discover new therapeutic uses for existing drugs while significantly reducing development time and cost. Traditional drug discovery methods rely heavily on manual laboratory experiments, expert analysis, and prolonged clinical trials, making the process slow, expensive, and limited in scalability. These approaches struggle to handle complex and high-dimensional biomedical data, leading to delayed insights and reduced efficiency. With the rapid growth of healthcare data, there is an increasing need for intelligent and automated systems that can efficiently analyze drug characteristics and predict alternative therapeutic applications. Additionally, conventional systems often lack transparency and strong security mechanisms, making clinical data vulnerable to tampering and reducing trust in research outcomes. To address these challenges, the proposed framework integrates Machine Learning (ML), Deep Learning (DL), and Blockchain technologies to develop a secure and intelligent drug repurposing system. The framework employs Random Forest (RF) as a baseline model and a Two-Dimensional Convolutional Neural Network (CNN2D) as an advanced model to improve prediction accuracy. The CNN2D effectively captures complex feature patterns in structured drug data, enabling precise identification of potential new disease treatments. Furthermore, Web3-based Blockchain technology ensures secure storage of user data, clinical interactions, and experimental records by providing immutability, transparency, and data integrity. By combining Artificial Intelligence (AI)-driven analytics with Blockchain-based security, the system enhances prediction performance, automates decision-making, and ensures reliable data management, offering a scalable and efficient solution for accelerating drug discovery and supporting healthcare innovation.

Key words: Healthcare Systems, Medical Data Processing, Drug Reprofiling, Machine learning, Deep Learning

1.INTRODUCTION

The discovery and development of pharmaceutical drugs is an inherently intricate and high-cost process that demands extensive scientific investigation, rigorous testing, and long-term investment. The complete lifecycle—from early-stage identification of promising chemical entities to final regulatory clearance—typically spans more than a decade, often ranging between 10 and 15 years. This extended timeframe reflects the multiple phases involved, including target identification, preclinical studies, clinical trials, and regulatory review, each of which introduces its own technical and operational complexities. Financially, the cost of developing a single drug frequently exceeds USD 1 billion, driven by expenses related to laboratory research, large-scale clinical trials, infrastructure, and compliance with strict regulatory standards [1]. Alongside economic challenges, the process is burdened by significant scientific uncertainties and logistical constraints, which can lead to repeated delays and increased attrition rates.

A major limitation in the drug development pipeline is the high failure rate observed during clinical trial phases. Many candidate compounds fail to progress due to insufficient efficacy or unexpected

safety concerns when tested in humans. Despite advancements in optimizing physicochemical and pharmacokinetic properties, translating laboratory success into clinically viable treatments remains a persistent challenge. These failures not only increase development costs but also highlight gaps in understanding complex biological systems and disease mechanisms [2]. The situation becomes even more challenging in the case of Cancer, a disease known for its remarkable genetic diversity and molecular complexity. Variations in gene expression, mutations, and signaling pathways across different cancer types—and even among patients with the same diagnosis—make it extremely difficult to design universally effective therapies. This heterogeneity complicates drug targeting, reduces treatment predictability, and necessitates highly personalized therapeutic approaches [3].

Considering the global burden and severity of cancer, there is an urgent demand for innovative strategies that can enhance the efficiency and precision of drug discovery. Traditional approaches alone are no longer sufficient to meet these growing demands, especially when dealing with large-scale biological data and complex disease interactions. As a result, advanced computational techniques are increasingly being incorporated into the drug discovery framework. In particular, Artificial Intelligence has emerged as a transformative technology with the potential to significantly improve various stages of the pipeline. By utilizing methods such as machine learning, Recurrent Neural Networks, deep neural networks, and evolutionary optimization techniques, AI enables the efficient analysis of vast biomedical datasets, identification of novel drug targets, prediction of molecular interactions, and optimization of compound structures. These capabilities allow researchers to make data-driven decisions, reduce experimental redundancy, and accelerate the identification of viable drug candidates. Ultimately, the integration of AI into drug discovery offers a promising pathway toward reducing development time, lowering costs, and improving the success rates of new therapeutic interventions [4].

Problem Definition

Existing drug repurposing methodologies suffer from issues like high latency, data silos, manual intervention, and lack of secure data handling. Moreover, the absence of automation and traceable validation hampers the reliability and reproducibility of research findings. The problem lies in the inefficient use of clinical and biochemical data and the vulnerability of centralized storage systems that pose a threat to data integrity and collaboration transparency.

2. RELATED WORK

2.1 AI-Driven Drug Repurposing and Discovery Trends

Recent advancements in computational healthcare have positioned drug repurposing as a practical and efficient alternative to traditional drug discovery. Melissa A. Haendel et al. [6] emphasized that repurposing existing drugs significantly reduces both development time and financial burden, particularly in the context of rare diseases where conventional approaches are often not viable. Their work demonstrated how AI can extract meaningful patterns from large-scale biomedical datasets to identify new therapeutic uses. Similarly, Cortial et al. [14] highlighted that although research in rare disease repurposing is still limited, the integration of AI is rapidly gaining momentum due to its ability to uncover hidden drug–disease relationships through data-driven approaches.

In parallel, Pandey et al. [10] showed that AI-based systems enhance virtual screening processes by efficiently analyzing molecular structures and predicting drug–target interactions. Their findings suggest that such systems not only accelerate candidate identification but also improve decision-making in early-stage drug design. These studies collectively indicate a paradigm shift toward data-centric methodologies in pharmaceutical research.

2.2 ML Techniques for Molecular Representation and Prediction

The application of ML has significantly improved the ability to process and interpret chemical and biological data. Patne et al. [7] focused on the use of SMILES-based representations, which convert complex molecular structures into linear textual formats that can be easily processed by computational models. By integrating ML with NLP techniques, their work enhanced key tasks such as lead identification, molecular interaction prediction, and virtual screening. This approach reduces dependency on time-consuming laboratory experiments while maintaining high predictive efficiency.

Further advancements are observed in the work of Qi et al. [8], who explored Transformer-based architectures for modeling molecular relationships. These models, originally designed for sequence-based tasks, demonstrated strong capability in learning structural dependencies within chemical compounds. However, their study also pointed out practical limitations, including the need for large datasets and high computational resources, which may restrict real-world applicability in resource-constrained environments.

2.3 Biological and Pharmacological Insights in Drug Effectiveness

Understanding the biological mechanisms of drug action remains a critical component in validating computational predictions. Ribeiro et al. [9] examined how specific compounds influence cancer-related processes such as cell proliferation, oxidative stress, and tumor microenvironment interactions. Their analysis provided valuable insights into how drugs affect signaling pathways and immune responses. Despite promising experimental outcomes, they emphasized that computational predictions must be supported by clinical validation to ensure reliability and safety.

2.4 Benefits and Limitations of AI in Pharmaceutical Systems

While AI has demonstrated substantial improvements in efficiency and predictive accuracy, several challenges persist. Kant et al. [11] discussed how AI enhances multiple stages of the drug development pipeline, including target identification, lead optimization, and toxicity prediction. However, they identified critical issues such as poor data quality, lack of interpretability, and regulatory barriers that limit large-scale adoption.

Similarly, Suri et al. [12] highlighted both the opportunities and constraints associated with AI integration. Their study pointed out that although AI enables faster discovery and supports personalized medicine, challenges such as data bias, privacy concerns, and biological complexity remain significant. In addition, Ali et al. [13] emphasized that DL models, despite their high performance, often suffer from issues related to generalization, data dependency, and computational cost.

3. PROPOSED SYSTEM

The proposed workflow integrates Artificial Intelligence with Blockchain to establish a secure and intelligent system for drug repurposing, enabling the discovery of new therapeutic uses for existing drugs, as illustrated in Fig. 2. The process begins with user registration and login through a secure interface, where user credentials and identity-related data are immutably stored on the blockchain to ensure transparency, integrity, and resistance to tampering. Once authenticated, users can upload a drug repurposing dataset containing critical biomedical information such as chemical composition, gene expression profiles, associated diseases, and symptom mappings. The system then performs a comprehensive preprocessing pipeline in which the uploaded data is cleaned to remove inconsistencies, missing values are handled, and the dataset is structured into a machine learning–ready format. Subsequently, the data is split into training and testing subsets using an 80:20 ratio to enable reliable model evaluation and generalization. Advanced AI models, specifically CNN2D and RF, are trained on the processed dataset to capture both spatial feature relationships and statistical patterns between drug compounds and disease characteristics. During training, the models learn complex associations that may

not be easily identifiable through traditional analytical approaches. After training, both models are rigorously evaluated using key performance metrics such as accuracy, precision, recall, and F1-score to ensure balanced and reliable prediction performance, with CNN generally demonstrating superior accuracy due to its ability to extract deep feature representations. Once the models are validated, the system allows users to input new drug-related data, which is processed and fed into the trained models to predict potential alternative diseases that the drug could effectively treat, thereby identifying novel repurposing opportunities. In addition to prediction, the system incorporates a collaborative research component where users can share trial results, experimental findings, and discussion insights with other researchers through a secure platform.

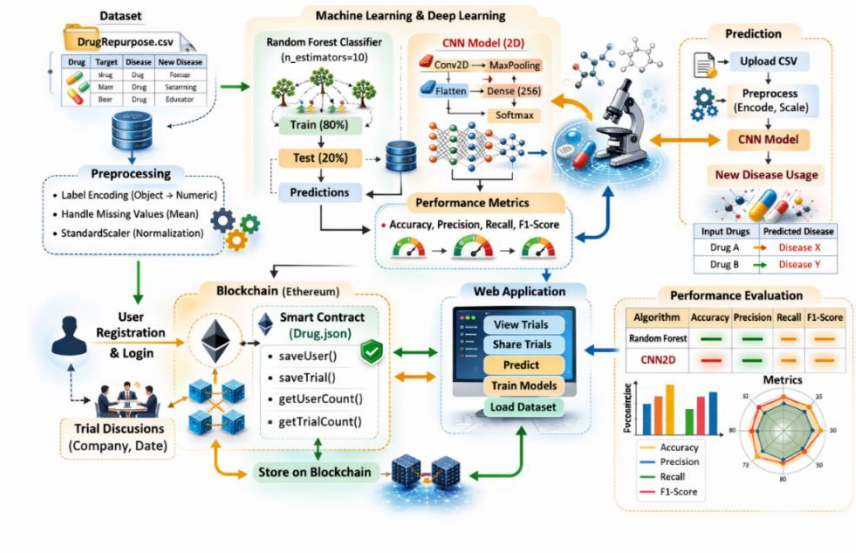

Fig. 2: proposed workflow.

These shared records are stored on the blockchain using smart contracts developed in Solidity, ensuring immutability, traceability, and secure knowledge exchange. Furthermore, all users can access and review these blockchain-stored trial records to gain insights, validate findings, and support collaborative decision-making. This integrated workflow combines secure data management, advanced AI-driven prediction, and decentralized collaboration to create a robust, transparent, and efficient drug repurposing ecosystem.

3.1 CNN2D

CNN learns hierarchical patterns from the input drug features (gene expressions, chemical descriptors, symptoms, etc.) using convolutional layers**,** activation functions**, and** pooling layers. The convolution layers extract spatial or sequential relationships between features. After sufficient pattern extraction and dimensionality reduction, flattened outputs are passed through dense layers for classification. This model is well-suited for structured biological data and gives high accuracy in predicting new drug uses.

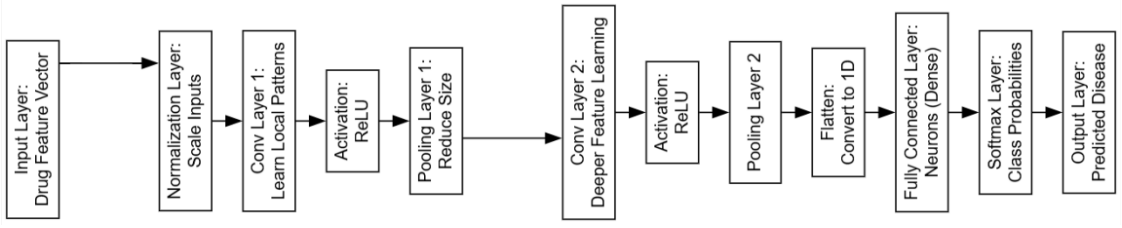

Fig. 3: Operational flow of CNN2D

- **Input Representation:** Drug data is converted into a 2D matrix (for example, gene expression as rows, chemical descriptors as columns), forming an **image-like representation.** Each pixel or cell corresponds to a normalized value of a biological feature.
- **Convolutional Layer(s):** Applies multiple filters (kernels) to the input matrix. Filters scan across the data, learning spatial feature relationships, such as how a drug's structural property aligns with certain gene expressions.
- **Activation (ReLU):** Introduces non-linearity, enabling the model to learn complex biological associations.
- **Pooling Layer:** Downsamples the data by taking maximum or average values in regions (e.g., 2x2 blocks). Reduces computational complexity and overfitting.
- **Deeper Layers:** Additional convolution and pooling layers are added to learn higher-order features. These may detect deeper biological relationships, such as how a chemical family interacts with specific disease pathways.
- **Flattening:** The final pooled matrix is converted into a 1D feature vector.
- **Fully Connected Layers:** Dense layers process the vector and output class probabilities for disease categories. The Softmax layer gives the likelihood that a drug can treat each disease.
- **Prediction:** The highest probability class is selected indicating the **most likely** repurposable **disease** for the input drug.

3.2 SMART CONTRACTS

The smart contract in this system as shown in Fig. 3 acts as a decentralized backend that securely stores all user registrations and clinical trial discussions. Instead of relying on a traditional database, the application sends data directly to the Ethereum network via Web3, ensuring that every transaction becomes tamper-proof and verifiable. When the Django application loads, it automatically connects to the blockchain, retrieves the ABI of the deployed contract, and makes it possible to call functions for storing or fetching information. As a result, every user entry, trial submission, or retrieval request is executed as a blockchain function call, creating a trusted and transparent environment for managing sensitive pharmaceutical records.

Establishing Blockchain Connection: The system begins by connecting to a local Ethereum node using Web3 and an HTTP provider. It sets the default account so the application knows which address will sign transactions. This establishes the communication channel necessary for sending and receiving blockchain data.

Loading the Smart Contract ABI and Address: The application loads the compiled contract JSON file to extract the ABI, which describes all available functions. It then loads the deployed contract address, allowing Web3 to link to the correct on-chain instance. This process ensures the Django system can interact with the exact version of the contract already deployed.

Fetching Data Stored in the Contract: Functions like getUserCount(), getUsername(), and getTrialCount() are called to read blockchain data. These calls do not require gas because reading data is free on the Ethereum network. The application loops through returned values and stores them locally so the interface can display user lists and trial discussions

Storing User Registration Data on the Blockchain: When a user registers, the system gathers inputs such as username, password, phone, and email. A transaction is created and sent to saveUser(), which

records the data permanently in the smart contract. The blockchain validates and mines the transaction, after which a receipt is returned confirming successful storage.

Storing Clinical Trial Discussions: Whenever a user submits trial details, the application calls saveTrial(), passing the username, discussion text, company, and date. This data is written to the blockchain as a new entry, ensuring it cannot be modified later. A confirmation receipt is generated, guaranteeing that the trial discussion is securely and immutably recorded.

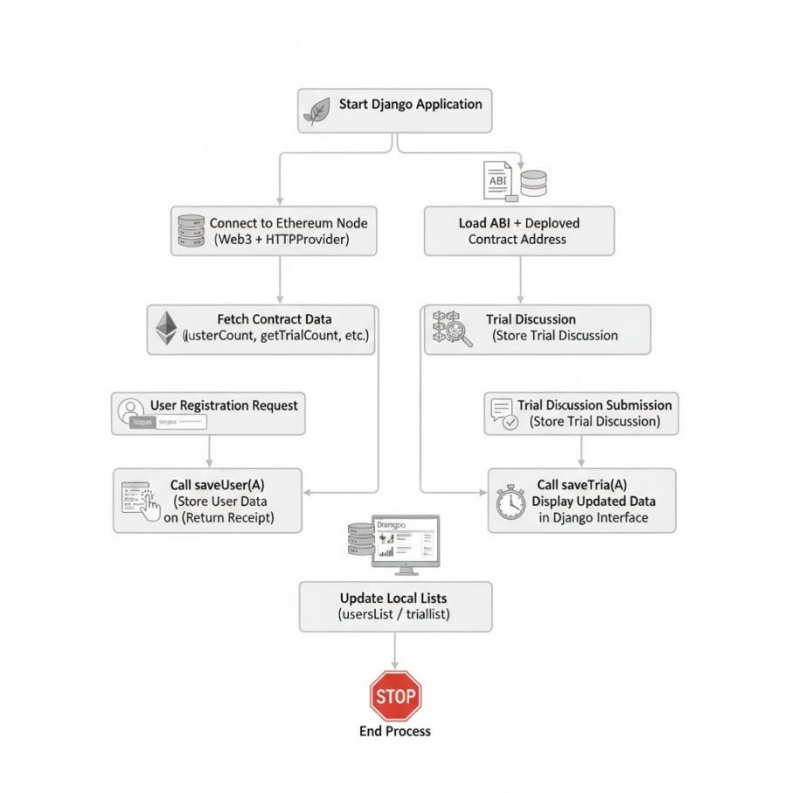

Fig. 3: Smart contract internal operation flow diagram.

Updating Local Lists After Blockchain Storage: After each successful transaction, the Django code appends the new user or trial data into local Python lists. This helps the frontend immediately display updated information without needing a full blockchain reload. The blockchain remains the authoritative source, while local lists enhance system responsiveness.

Displaying Data to the User Interface: The application reads the structured values retrieved from the smart contract and generates HTML tables dynamically. Trial history, user information, and other blockchain records are presented clearly in the interface. This makes the blockchain-powered backend appear seamless to end-users while maintaining maximum security.

4. RESULTS ANALYSIS

Algorithm Name	Accuracy	Precision	Recall	FSCORE
Random Forest	94.154	93.341	94.336	92.695
CNN2D	93.737	92.081	93.475	91.63

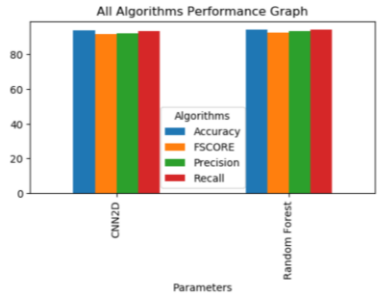

Fig. 4: Model Training and Performance Metrics Screen

Fig. 4 illustrates the model training and performance evaluation stage of the proposed system. This screen presents a comparative analysis of CNN2D and RF models using standard evaluation metrics such as accuracy, precision, recall, and F1-score. The performance comparison highlights the effectiveness of feature learning and classification capability of each model. Graphical representation supports quantitative assessment and model selection. This stage validates the predictive reliability of the trained models before deployment.

Fig. 5 depicts the clinical trial sharing module integrated with blockchain technology. This interface allows authenticated users to submit trial-related information for secure and immutable storage. Each submission is recorded as a blockchain transaction, ensuring transparency and traceability. The decentralized ledger prevents tampering and supports collaborative research. This figure highlights the role of blockchain in maintaining trust in shared clinical data.

DrugID	Formula	DrugName	Target	Disease	New Disease Usage
DB13855	C12H9N3O5	Nifuroxazide	Bacterial enzymes	Diarrhoea	Can be used for New Disease = Pulmonary fibrosis
DB00746	C25H48N6O8	Deferoxamine	Iron ions	Acute iron or aluminum toxicity	Can be used for New Disease = Pompe disease
DB14480	C5H10N2O2S	Acetylcysteine amide	glutathione	Acetaminophen toxicity	Can be used for New Disease = Dengue
DB00571	C16H21NO2	Propranolol	ADRB1	Migraine	Can be used for New Disease = Colorectal cancer
DB01136	C24H26N2O4	Carvedilol	Adrenergic receptors	Congestive heart failure; Hypertension; Depression	Can be used for New Disease = Epileptic encephalopathy
DB00675	C26H29NO	Tamoxifen	ESR1	Breast cancer	Can be used for New Disease = Immune-mediated diseases
DB01069	C17H20N2S	Promethazine	HRH1	Nausea	Can be used for New Disease = Colorectal cancer
DB00786	C15H29N3O5	Marimastat	MMP1; MMP2; MMP7	Pancreatic cancer; Lung cancer; Tissue repair disorder; Tumor development; Inflammatory disorder	Can be used for New Disease = Venom of the western diamondback rattlesnake, Crotalus atrox
DB06145	C43H74N2O14	Spiramycin	Gram-positive bacteria	Bacterial infection	Can be used for New Disease = Anti-inflammatory
DB00571	C16H21NO2	Propranolol	ADRB1	Migraine	Can be used for New Disease =

Fig. 5: Predicting Test Data Screen

5. CONCLUSION

The developed blockchain-enabled AI framework demonstrates an effective integration of ML, DL, and decentralized storage for predictive drug repurposing. The system efficiently preprocesses large-scale datasets, encodes complex biomedical features, and utilizes RF and CNN2D models to predict alternative drug applications. Experimental evaluation shows that RF slightly outperforms CNN2D in terms of accuracy, precision, recall, and F1-score, indicating better stability for this dataset. Performance gains are achieved through proper scaling, balanced data splitting, and effective handling of categorical features. The CNN benefits from GPU-based optimization, ensuring faster training and stable convergence. Blockchain integration ensures secure, transparent, and tamper-resistant storage of user and trial data. Smart contract validation using Web3.py confirms data integrity and seamless interaction between AI models and the decentralized system. Overall, the framework provides a reliable, secure, and user-friendly solution for drug repurposing through a structured Django interface.

REFERENCES

[1] Hughes, J.P.; Rees, S.; Kalindjian, S.B.; Philpott, K.L. Principles of early drug discovery. Br. J. Pharmacol. 2011, 162, 1239–1249.

[2] Waring, M.J.; Arrowsmith, J.; Leach, A.R.; Leeson, P.D.; Mandrell, S.; Owen, R.M.; Pairaudeau, G.; Pennie, W.D.; Pickett, S.D.; Wang, J.; et al. An analysis of the attrition of drug candidates from four major pharmaceutical companies. Nat. Rev. Drug Discov. 2015, 14, 475–486.

[3] Vasan, N.; Baselga, J.; Hyman, D.M. A view on drug resistance in cancer. Nature 2019, 575, 299–309.

[4] Garg, P.; Singhal, G.; Kulkarni, P.; Horne, D.; Salgia, R.; Singhal, S.S. Artificial Intelligence-Driven Computational Approaches in the Development of Anticancer Drugs. Cancers 2024, 16, 3884.

[5] Wang, L.; Song, Y.; Wang, H.; Zhang, X.; Wang, M.; He, J.; Li, S.; Zhang, L.; Li, K.; Cao, L. Advances of Artificial Intelligence in Anti-Cancer Drug Design: A Review of the Past Decade. Pharmaceuticals 2023, 16, 253.

[6] Haendel M, Vasilevsky N, Unni D, Bologa C, Harris N, Rehm H, Hamosh A, Baynam G, Groza T, McMurry J, Dawkins H, Rath A, Thaxton C, Bocci G, Joachimiak MP, Köhler S, Robinson PN, Mungall C, Oprea TI. How many rare diseases are there? Nat Rev Drug Discov. 2020 Feb;19(2):77-78. doi: 10.1038/d41573-019-00180-y,PMID: 32020066; PMCID: PMC7771654.

[7] Patne AY, Dhulipala SM, Lawless W, Prakash S, Mohapatra SS, Mohapatra S. Drug Discovery in the Age of Artificial Intelligence: Transformative Target-Based Approaches. *International Journal of Molecular Sciences*. 2024; 25(22):12233. https://doi.org/10.3390/ijms252212233

[8] Qi X, Zhao Y, Qi Z, Hou S, Chen J. Machine Learning Empowering Drug Discovery: Applications, Opportunities and Challenges. *Molecules*. 2024; 29(4):903. https://doi.org/10.3390/molecules29040903

[9] Ribeiro E, Vale N. Repurposing Terbutaline and Milrinone for Cancer Therapy: A Comprehensive Review. *Future Pharmacology*. 2025; 5(3):38. https://doi.org/10.3390/futurepharmacol5030038

[10] Pandey, B.S., Durgapal, S., Kumar, P. et al. Unraveling the artificial intelligence role in drug discovery and pharmaceutical product design: an opportunity and challenges. Discov Artif Intell 5, 72 (2025). https://doi.org/10.1007/s44163-025-00330-9

[11] Kant, S., Deepika & Roy, S. Artificial intelligence in drug discovery and development: transforming challenges into opportunities. Discov. Pharm. Sci. **1**, 7 (2025). https://doi.org/10.1007/s44395-025-00007-3

[12] Suri, G.S., Kaur, G. & Shinde, D. Beyond boundaries: exploring the transformative power of AI in pharmaceuticals. Discov Artif Intell **4**, 82 (2024). https://doi.org/10.1007/s44163-024-00192-7

[13] Ali, N., Hanif, N., Khan, H.A. et al. Deep learning and artificial intelligence for drug discovery, application, challenge, and future perspectives. Discov Appl Sci **7**, 533 (2025). https://doi.org/10.1007/s42452-025-06991-6

[14] Cortial L, Montero V, Tourlet S, Del Bano J, Blin O. Artificial intelligence in drug repurposing for rare diseases: a mini-review. Front Med (Lausanne). 2024 May 22;11:1404338. doi: 10.3389/fmed.2024.1404338,PMID: 38841574; PMCID: PMC11150798.