AI in Drug Discovery and Development

Prof. Rajnish Kumar
Dept. of Pharmaceutical Engineering and Technology
IIT-(BHU), Varanasi

Week-11
Lecture-55

Welcome to the course "AI in Drug Discovery and Development." In this session, we will talk about the future outlook of explainable AI and other emerging technologies in drug discovery. So, until now we have seen that AI is, you know, extremely useful in drug discovery and development, especially in expediting it as well as reducing its cost. However, it is still emerging, and there are new technologies coming up. So, in this session, we will talk about especially explainable AI, as well as other emerging technologies that are useful in drug discovery and development. So, by the end of this lecture, you will be able to recognize current limitations in AI-driven drug discovery that motivate the need for next-generation technologies.

Understand how explainable AI improves transparency, trust, and regulatory compliance. Explore the role of foundation models, quantum computing, hybrid AI-quantum approaches, and digital twins in modern drug discovery. as well as identifying key applications of emerging technologies, including de novo drug design, ADMET prediction, and personalized medicine simulations. So, we have talked a lot about those challenges in drug discovery and utilizing AI for drug discovery and development in earlier sessions, but specifically about the explainability of those black box models.

So, this hinders the interpretability, reducing trust and adoption. So, this is one of the biggest challenges of those models due to the black box nature as well as poor generalization. which is due to the bias or limited training dataset that affects the performance on diverse datasets as well. And then there were regulatory barriers that demand explainable and transparent AI systems for healthcare approval and low clinical translation of AI predictions, which highlight the need for robust validation. So, these are some of the challenges that derive from the future needs as well as the future technologies that are coming up to resolve the challenges.

So, we compare the white box, grey box, or black box nature of the models. For example, if you talk about a white box model, So, you can inspect the inner logic of an ML model as well as understand its decisions. So, that is how it is having a higher explainability and interpretability. However, it has low accuracy, and due to its low accuracy, it is not used in daily life applications, and this is, you know, if you go from like bottom to up, bottom up. So, you can see that self-interpretability is increasing.

However, at the bottom, the interpretability is very low. So, these are neither interpretable nor explainable. So, then you have the grey box models where these are partially analyzed, the internal workings of a model are understood, and you can interpret to some degree with significant accuracy. They can be used in critical applications if designed carefully, and then we come to the black box nature of these models, especially those deep learning models. Where you could not inspect the inner logic of the model and understand its decisions.

However, they have higher accuracy but lower explainability and interpretability. So, due to a non-explainable decision, it is not practical to use in critical applications. However, with the help of these XAI explainable AI models, we can get rid of, you know, this challenge. So, the explainability or interpretation it is revealing the decision making mechanism and then what it does is it build the user trust being fair and ethical where you can verify the prediction. And then interpretability is disclosing the internal workings of the model.

So, it is through this that you understand the intrinsic property of the model that enhances the model's transparency, and ultimately it leads to the trustworthiness of the model. So, explainable AI is actually a solution to the limited interpretability and explainability of the models. So, XAI refers to techniques and models that allow humans to understand, trust, and effectively manage AI decision-making processes. So, traditional AI models, as we are talking about DL models, often function as black boxes providing accurate predictions but without clear reasoning. So, these XAI models they aim to make these complex models transparent, interpretable, and enabling user to see how and why decisions are made.

So, this is, you know, a knowledge graph that shows the interconnected potential uses of the explainability concepts. So, ultimately, it is leading to the trustworthiness and different components that are merging and bringing the trustworthiness. For example, responsibility, which you know requires fairness, is further needed by the stability and satisfaction that require robustness. It extends to explainability, contributes to interpretability, fosters interactivity, and increases transparency. So, likewise, all these parameters are interconnected to each other somehow, and they all contribute to the trustworthiness of the model.

So, why is XAI crucial in drug discovery? Because it builds trust with researchers and clinicians, while black box AI systems hinder adoption due to their lack of transparency. and interpretability enhance user confidence and encourage real-world application Furthermore, the regulatory and ethical requirements set by healthcare regulators like the FDA or AMA mandate transparency in AI-assisted decision-making, and XAI supports

auditability and compliance with the ethical guidelines as well. So, they also enhances the scientific understanding by offering insight into model behavior, helping researchers identify novel drug mechanisms, targets or biomarkers and further facilitates the hypothesis generation and validation. And it ultimately improves the translational success, where the transparent models help explain why a drug candidate is selected. Reducing false positives as well as enhancing the reproducibility and alignment with the clinical outcomes.

So some of the key XAI technologies used in drug discovery are, for example, the feature importance technique, which ranks input features, such as molecular properties, by their influence on the predictions. And then they can be used to identify, for example, the key pharmacophores and the ADME properties. And then you have the SHAP method, which is Shapley additive explanations, which quantifies each feature's contribution to the prediction. So, by using the SHAP method, we can understand which molecular attribute drive the efficacy and which attribute is responsible for, you know, that affecting that endpoint. And then you have LIME, which stands for Local Interpretable Model-agnostic Explanations, and it provides local explanations for individual predictions.

So, it analyzes the outliers in unexpected predictions. And then you have the attention mechanisms, which are being used in, you know, deep learning, for example, transformers, to focus on important input regions. So, it can be used to reveal the binding sites in the protein-ligand complexes. So in medicinal chemistry, XAI can explain decisions in drug design by prioritizing compounds. We can identify molecular features linked to bioavailability, binding affinity, toxicity, or permeability.

So all those sorts of, you know, ADMET properties or efficacy, you know, bioactivities. So we can point out which part of those molecules is responsible for that feature, and then ultimately we can optimize those molecules based on that information. We can help us understand the SARs revealing how structural changes affect the activity to guide the optimization. As well as visualizing the key features, tools like saliency maps highlight functional groups driving toxicity or efficacy. So, XAI can be used in toxicity prediction, where toxicity prediction is crucial in drug development.

XAI enhances the process by using tools like DeepTox or DeepPurpose to predict toxicity from molecular structures. It applies a sharp explanation to pinpoint toxic substructures or the patterns, and this not only boosts the prediction accuracy but also guides the design of safer drug candidates. However, there are always, you know, limitations to everything, and explainable AI also comes with some limitations; these include, for example, the trade-off with model performance. So, the simpler interpretable models like decision trees often sacrifice accuracy compared to the complex black-box models like deep neural nets, and then the post hoc explanations can be misleading. Many XI methods explain predictions

after the fact, which might not truly reflect the model's internal logic as well.

A lack of standardization means there is no universal framework for evaluating or validating explanations, making it hard to compare across methods or domains. And then there are scalability issues, such as generating explanations for large datasets or complex models, which can be computationally expensive. As well as the domain expertise requirement, understanding and interpreting the output of XAI tools often requires deep domain knowledge, especially in fields like medicinal chemistry. Overinterpretation risk, so users might over-rely on or misinterpret the explanations, leading to false confidence in the model's decision. As well as limited causality insights, mostly XAI tools show correlation, not causation, which may not be sufficient for critical decisions in healthcare or drug development.

Another emerging technology is the foundation model. So, the foundation models are large-scale AI systems pre-trained on vast and diverse data sets using self-supervised learning techniques. And the key feature is that you can train them once, and you can apply them many times. So, these are pre-trained on general data and then fine-tuned or directly used for specific applications. So, some of the core characteristics of these foundational models are scale and generality.

So, these are trained zones, you know, terabytes of multi-modal data; it could be like text, sequence structure, etc. And they are capable of capturing rich representations across the domains. And another characteristic is the transferability, which can be adapted to diverse tasks like classification, generation, prediction, and translation. And then you have the few short and zero short learning situations where it is effective even with minimal task-specific examples as well. And then multimodality, especially in the newer models.

They process multiple data types, such as text, molecules, protein sequences, or even 3D structures. So, what is another key characteristic of the foundation models? So, some of the foundation models being used in drug discovery are like GPT, which is a language model, and it can be, you know, pre-trained on natural language adapted for text mining and hypothesis generation. And then you have AlphaFold, which is applicable in structural biology and predicts the three-dimensional protein structure from the amino acid sequences. And then you have MolBERT, which is a cheminformatic tool. So, the transformer model is pre-trained on SMILES for molecular property prediction and generation.

And then you have the ESM fold, which again uses large transformers for accurate protein structure prediction and is a good alternative to AlphaFold. And then you have ProtBERT or ProTrans, which are protein language models pre-trained on millions of protein

sequences for tasks like functional annotation and variant effect prediction. So, there are multimodal foundation models that can take multiple inputs, and these inputs could be text, for example, in the form of scientific literature or experimental protocols. or molecular representations in the form of SMILES or InChI keys, or protein sequences in the form of amino acid chains and 3D structures in the form of atomic coordinates, or folding data like in the form of PDB or MMCIF. And some examples of multimodal models are BioGPT, which is a domain-specific language model for biomedical question and answer or literature mining and knowledge discovery.

Galactica, which is trained on scientific text for tasks like citation, prediction, reasoning, and summarization, is designed for various applications. ProGPT2 is a transformer-based model for generating de novo protein sequences with realistic biological properties. Talking about the applications of foundation models in drug discovery, this can be used in, you know, de novo drug design, where they can generate novel compounds with desired biological properties. And then they can be used in target identification validation, where they can leverage pre-trained protein embeddings to identify the disease-relevant proteins. And then they can use it for the ADMET prediction.

The models like MOLBERT can be used for predicting ADMET properties. And then, these models can also be used for virtual screening, where they can accelerate high-throughput screening using molecular embeddings from the pre-trained models. And then you can also do the protein structure and functional prediction, which improves the efficiency and accuracy of the drug target interaction studies. And then it can be used for literature mining and hypothesis generation. The models like BioGPT can be used to uncover hidden relationships in the biomedical literature.

Again, some of the challenges with these multimodal models, or the foundation models, require a lot of computational power. They are computationally quite expensive. So, training and fine tuning those models requires significant resources. And then interpretability can be another issue. So, the output may lack a clear biological rationale.

And then there is a lot of data bias, as well as the models can inherit bias from the training data. And then there is regulatory uncertainty as well. So, using clinical decision-making is still under scrutiny because regulatory agencies, you know, want to make sure that they are good at doing that task. Another emerging technology is quantum computing. So, quantum computing leverages principles of quantum mechanics such as superposition, entanglement, and quantum tunneling to perform computations far beyond the capability of classical computers.

So, instead of, you know, the classical bits, qubits are used; quantum computers have

massively parallel processing potential, and these are ideal for solving complex optimization simulation problems. So, why do we pursue quantum computing for drug discovery? Traditional drug discovery is computationally quite expensive when we talk about molecular dynamics or even molecular docking. So, they actually take a lot of computational time, and they are limited in simulating large quantum-level systems if you want to simulate an entire cell or a whole viral particle. So, I think using molecular dynamics is really, really difficult to simulate large systems. So, with the help of quantum computing, it can be possible.

So, quantum computing offers accurate quantum mechanical simulations of molecules as well as efficient exploration of chemical space. as well as they have potential to predict binding affinities, reaction pathways and protein folding with the high fidelity. So, some of the key applications could be in molecular simulation, where they can help model the electronic structure interactions at the atomic scale. In protein folding, where they can predict the 3D structure with quantum-inspired optimization. They can be used for ligand-receptor binding studies, where they can improve the docking by quantum-enhanced scoring functions.

They can be used for de novo drug design, where they can optimize the molecular properties across vast chemical space. And then they can be used for, you know, reaction mechanism discovery, where they can be used to simulate the reaction kinetics and the transition states. Okay, coming to another emerging tech, which is the hybrid AI plus quantum approaches for track discovery. So, these hybrid approaches combine classical AI techniques like ML and DL with quantum computing to leverage the strengths of both fields. So, AI can handle large-scale data analysis, pattern recognition, and predictive modeling, while quantum computing tackles computationally intensive quantum mechanical simulations such as protein folding and molecular interactions.

So, the combination results in faster, more accurate drug discovery pipelines. So, why do we have to have the hybrid, you know, AI plus quantum approaches? Because AI can be used for scalability, the AI models process and analyze large data sets, identify patterns, and optimize drug candidates efficiently. However, the quantum can be used for precision, where the quantum computer simulates quantum-level interactions. Protein-ligand binding interactions and the electronic structure of the molecules that are beyond the reach of classical simulation. They are used for their speed and accuracy; together, they reduce time and increase precision in key drug discovery tasks such as molecule design and protein structure prediction.

So, let's talk about the applications of hybrid approaches in drug discovery. So, this can be used in molecular simulations. The AI can be used to predict molecular properties like

binding affinity, and quantum computing can simulate molecular interactions at the electronic level. In design and optimization, AI models can suggest drug candidates from large datasets, while quantum computing defines interactions and energetics in molecular structures. In protein folding, AI techniques identify folding patterns and predict stability, while quantum computing calculates quantum states and predicts protein conformations more accurately.

In ligand-receptor binding, AI techniques analyze ligand-receptor databases to find promising interactions, and quantum computing simulates binding affinity with high precision at the atomic level. And then regarding the reaction pathways, AI can help in predicting the reaction outcomes from the historical data, while quantum computing simulates reaction mechanisms, transition states, and the energy profiles. So, some of you know the hybrid models along with the collaboration. So, Google started integrating quantum AIML, you know. So, combining quantum simulation with AI to accelerate molecular dynamics and drug discovery.

And then there is this IBM QIS kit and AI. The QIS kit integrates classical ML with quantum computing for more efficient drug candidate screening. And then Boehringer and Google are using hybrid approaches to model protein folding and enhance the drug discovery pipeline. And then Microsoft is using quantum and DL, employing hybrid quantum and classical algorithms to optimize molecular properties and predict drug target interactions. So, talking about the challenges with quantum computing. So, the quantum hardware limitation is one of the big challenges because the current quantum computers have limited qubit counts and a relatively large error rate as well.

And then the complexity of the hybrid system integrating quantum computing with AI models requires expertise in both fields and careful design of the hybrid algorithms. Then, scalability is another issue where hybrid approaches are still limited to small molecules and systems due to quantum computational constraints. Data complexity is another challenge where the high-dimensional quantum data needs to be processed effectively using AI models that require advanced techniques in data compression and model interpretability. Another emerging technique is the digital twins, which we talked about in, you know, clinical trials as well. So, a digital twin is a virtual model or simulation that mirrors a real-world entity or system.

So, in drug discovery, a digital twin refers to a virtual representation of a biological system. Molecules or even patients created using data from various sources, such as genomics, imaging, and clinical trials. The probability of a digital twin is dynamic in nature because it is continuously updated with real-time data. And then it is predictive as well, which can be used for simulation and prediction of future behaviors or outcomes. So, how do digital

twins work in drug discovery? It actually starts with data integration.

So, combining the genomic data, proteomics data, clinical records, and real-time biological data from the patients, lab experiments, and the sensor, a model is then built. So, building a virtual representation of biological systems at the molecular, cellular, tissue, and organ levels. And then doing the simulation prediction, where it uses the model to predict the drug responses, interactions, and potential side effects in individuals or populations. And finally, optimization where the model is continuously updated with new data and refines predictions for better decision-making in drug design and patient care. So, some of the applications of digital twins in drug discovery are that they can be used for personalized drug development, which can tailor drugs to individual genetic profiles and health conditions.

They can be used for toxicity prediction as well by simulating how drugs interact with human tissues and organs to predict the adverse effects. And then, for disease modeling, they can also be used; we can create models of disease progression and simulate drug efficacy in virtual patients. For preclinical testing, it can be used to simulate clinical trial outcomes in a virtual population before the physical trials. And the clinical trial optimization as well, where we can predict optimal dosing, patient recruitment, and treatment regimens for the clinical trials. So, one of the use cases of digital twins is in silico medicine, where they have in silico a tool that can be used for conducting in silico clinical trials.

Then it uses AI and digital twins to identify drug targets, optimize compounds, and predict the outcomes. And then you have another example, GNS Healthcare, which builds digital twin operations using AI to simulate treatment outcomes. It has predicted which therapies are likely to succeed for different patient subgroups. Some of the other emerging technologies are, like you know, the generative chemistry platforms we talked about, the re-invent, which is a reinforcement learning-based platform for molecular generation.

So, that is one of the things. Another is the Chemcrow, which is a really wonderful tool. So, it combines language models and chemistry tools for autonomous compound design. It is like a virtual assistant that designs molecules for you. So, you just ask it if it can create some structures of molecules having a molecular weight from this to this and having good solubility.

It can create those actually. So, it has that capability. And then there are these AI-enhanced molecular dynamic simulations where the integration of MD with reinforcement learning and active learning optimizes the simulation pathways in real time. So, the exemplary models are ANI, an accurate neural network potential. Which is a deep learning model

trained on quantum chemical data, and the MDCROW is another tool from the ChemCROW team. So, this is for automating the molecular dynamics workflows with a large language model. So, to MDCROW you can just tell ok I can you run like 100 nanosecond fully atomistic molecular dynamic simulation for a protein with the PDBID of 3EY7.

And then it can actually perform that, analyze the data, make a report, and get back to you. So, it is still under development. So, it is really good; however, there is some scope for improvement, and those people are working hard on it. Okay, coming to the summary, summarizing all this, those explainable tools, explainable AI, and emerging technologies explain. So AI is transforming drug discovery, but major gaps remain in interpretability, scalability, and clinical translation.

And XAI enables more transparent, trustworthy and regulatory aligned decision making. So, the formulation models they offer generalizable solutions across multiple tasks from compound generation to ADMET prediction. And the quantum computing and hybrid approaches, which combine AI and quantum methods, bring unprecedented precision to molecular simulations and optimization. And the digital twins enable dynamic, data-driven models of biological systems and patients for in silico testing. And the future innovations depend on the responsible integration of emerging technologies that are robust, scalable, and aligned with real-world needs.

In the end, I have an open question for you. Could the first inhuman trial one day occur entirely within a machine? Because of the development of all those, you know, fancy techniques, I believe we will have that time quite soon, actually. So, just think about it. And I have some suggestions for further reading where you can go through these papers to learn more about this topic. And with that, thank you.