THE USE OF AI AND ADVANCED COMPUTER SYSTEMS TO DEVELOP DRUGS AGAINST NEW EMERGING THREATS

Accelerating Drug Discovery And Repurposing With MIRA: A Machine Learning And Genai Approach

Methun Kamruzzaman Sandia National Labs **Kenneth Sale** Sandia National Labs **Kunal Porrey** Sandia National Labs

The robust maintenance of a continuous supply of pharmaceuticals, coupled with the capacity for developing novel therapeutics in response to emerging diseases, is vital for enhancing national healthcare resilience. Challenges such as supply chain disruptions, propriety formulations, drug resistance, and reliance on single-source raw materials or the products itself can hinder access to proper medical countermeasures. Machine learning (ML) aided drug discovery significantly reduces bottlenecks by enhancing both the speed and cost-effectiveness of the process. By using extensive datasets, ML algorithms identify potential drugs, predict interactions with biological targets, and optimize chemical properties for effective manufacturing. New drug development also aids in improving treatment efficacy and reducing development costs. There are still needs in developing "explainable" machine learning (XML) methods that enhances our understanding of structure-property relationships, aiding in the optimized design of effective and safer drugs by understanding model recommendations, identifying new drug targets, and predicting side effects. Additionally, we explore the potential of generative artificial intelligence (GenAI) in leveraging this knowledge to further enhance accelerated drug development.

We introduce MIRA (Machine Intelligence for Rapid Acceleration of Drug Discovery and Repurposing), an innovative model combining GenAI tools and XML for accelerated drug discovery. MIRA integrates state-of-the-art GenAI models capable of conditionally generating drug-like molecules. And utilizing publicly available databases, we've compiled a database of drug compounds and their properties, such as toxicity, solubility, and permeability. MIRA employs cheminformatics to predict the properties of both existing and synthetic structures, offering ranked alternatives for drug compounds for repurposing and elucidating structural-property relationships and designing new ones. This approach accelerates experimental testing and validation by in-sillico screening and reducing the experimental burden for validation. MIRA present the capability of applying generative models based on large language models (LSTM and GPT) for generating novel alternative drug-like molecules. We demonstrate the use of this capability on characterizing properties of drugs to pass blood brain barrier. This holistic approach paves a promising future in more accurate and informed drug discovery.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy’s National Nuclear Security Administration under contract DE-NA-0003525.