AI Driven Drug Discovery: Designing Cancer Inhibiting Molecules With Genetic Algorithm

Kim, Claire (School: Lamar Academy)

Globally, health issues such as antibiotic-resistant infections and cancer become increasingly severe; however, the extensive time and cost required for drug development impose a significant burden. To address these challenges, there is a growing effort to leverage machine learning generative algorithms and shorten the drug development cycle. However, current models often fail to guarantee the validity, synthesizability, and efficacy of the generated molecules. This research aims to design a Genetic Algorithm (GA)-based molecule generation model, OptiMolGA, that ensures molecular validity while maintaining efficiency, and to take a step toward the active implementation of AI models in cancer therapy. Unlike Deep Learning models, OptiMolGA uses evolutionary processes to diversify and refine solutions. These principles, such as mutation and crossover, are combined with a scoring function that incorporates Quantitative Estimate of Drug-Likeness (QED), Synthetic Accessibility (SA), and Vina Docking scores. Using the CrossDocked2020 dataset to compare the performance of OptiMolGA with other models, the model consistently refined and generated candidate molecules over more than 500 runs. Experimental results show that, on average, OptiMolGA outperforms state-of-the-art target-aware molecule generation Deep Learning models in all QED, SA, and Vina Docking scores. Subsequently, OptiMolGA was applied to generate potential inhibitors for crystal structures of various cancer-related proteins and mutations. The inhibitors generated by OptiMolGA demonstrated Vina Docking scores that were equal to or superior to those of active, existing inhibitors. Overall, the success of OptiMolGA highlights its tremendous potential to become an effective and innovative tool in drug discovery.

Awards Won: