

## Połączenie symulacji molekularnej i uczenia głębokiego w projektowaniu leków *de novo*

Odkrywanie leków jest długim i kosztownym procesem, który często kończy się porażką, a jego koszty rosną z każdym kolejnym rokiem. Szczególnie teraz, kiedy świat ciągle zmagają się z pandemią COVID-19, oczywiste staje się, że chemia medyczna skorzystałaby z wprowadzenia narzędzi redukujących koszty tego procesu i skracających drogę do odkrycia nowych skutecznych leków. Jednym ze sposobów, by to osiągnąć, jest zastosowanie mocy obliczeniowej do komputerowego rozwiązania zadań projektowania leków. W szczególności sztuczna inteligencja została już skutecznie użyta w projektowaniu leków wspomaganym komputerowo do rozwiązania wielu zadań, takich jak identyfikacja celu, wirtualny skrining i optymalizacja molekuł. Rozwiązania oparte na uczeniu maszynowym w dużym stopniu jednak zależą od jakości danych, które mogą być mylne w obszarze farmaceutyki. Firmy mają tendencję do raportowania tylko udanych związków chemicznych lub wcale nie publikują swoich odkryć, a zatem tylko mała część ogromnej przestrzeni chemicznej może być wyeksplorowana przez modele uczenia maszynowego. Prowadzi to także do częstej krytyki modeli za brak nowości proponowanych molekuł lub tworzenie nierealistycznych związków, które nie mogą być zsyntetyzowane.

Wraz ze zwiększającą się ilością danych chemicznych przychodzi potrzeba metod, które prześwietliłyby dostępne biblioteki związków chemicznych w celu znalezienia molekuł dobrze oddziałujących z wybranym celem. Do tego zadania zostały zaproponowane programy dokingu molekularnego, które sprawiają, że mozolny proces odkrywania leków staje się wykonalny. Programy dokingowe symulują optymalne ułożenie molekuły w środku kieszeni wiążącej wybranego celu. W szczególności chemicy używają programu dokingowego do badania interakcji pomiędzy małymi molekułami a białkami i podejmowania decyzji z większą wnikliwością podczas projektowania nowych leków. Programy dokujące również ułatwiają poszukiwanie potencjalnie aktywnych związków w dużych bibliotekach związków chemicznych, co zwane jest wirtualnym skriningiem. Ostatnio uczenie głębokie zostało użyte by przyspieszyć prześwietlenie ogromnych zbiorów chemicznych poprzez przewidywanie wyniku programu dokującego przez szybszą od tradycyjnego dokingu sieć neuronową.

Głównym celem projektu jest *eksploracja potencjału, jaki niesie doking molekularny dla projektowania leków de novo ze szczególnym uwzględnieniem jakości modeli opartych na uczeniu głębokim*. Mówiąc dokładniej, proponujemy nową wysokopoziomową metodologię łączenia symulacji z dowolnym modelem generatywnym, służącym do projektowania leków *de novo* lub optymalizacji molekuł. Zostanie zaimplementowanych kilka modeli generatywnych i będą one tworzyć nowe związki chemiczne o poprawionych własnościach. Zeby śledzić te własności, zostaną również zaimplementowane modele predykcyjne. Następnie zbadamy różne metody wprowadzania do modelu informacji z symulacji, np. trenowanie modelu na danych symulowanych lub taka optymalizacja modelu, by generował molekuły osiągające dobre wyniki w symulacji.

Jako efekt końcowy zostanie zaimplementowany model głęboki, który generuje cząsteczki zgodnie z opisaną metodologią. Stworzony model będzie użyteczny w projektowaniu leków *de novo*, czyli tworzeniu nowych potencjalnych leków, bazując na zadanym celu biologicznym, ale także w optymalizacji struktury wiodącej, przez co rozumie się optymalizację istniejących związków (związków wiodących) w celu poprawienia ich własności fizykochemicznych i zidentyfikowania potencjalnych leków.