

Spektroskopia NMR jako język molekularny: podejście oparte na uczeniu głębokim do przewidywania bioaktywności

Sztuczna inteligencja (AI) wkracza do świata chemii nie tylko jako narzędzie przyspieszające analizę danych, ale jako technologia, która zmienia sposób, w jaki patrzymy na materię. W tradycyjnych modelach komputerowych cząsteczki przedstawiane są za pomocą tzw. „odcisków palców” – uproszczonych, binarnych wektorów, które kodują jedynie podstawowe informacje o strukturze związku. Taki zapis, choć wygodny, gubi wiele istotnych szczegółów – jak rzeczywisty kształt cząsteczki, rozkład elektronów czy wpływ otoczenia biologicznego.

W tym projekcie proponujemy coś więcej niż tylko „matematyczne skróty”. Zamiast uczyć algorytmy rozpoznawania uproszczonych symboli chemicznych, dajemy im narzędzia, które pozwalają „widzieć” i „słyszeć” cząsteczkę niemal tak jak człowiek – poprzez **widma spektroskopowe**, a konkretnie przez **wirtualne widma NMR**, generowane na podstawie struktury chemicznej.

Widma NMR to jedno z najważniejszych narzędzi chemii eksperymentalnej – pozwalają zobaczyć, jakie atomy znajdują się w cząsteczce, jak są połączone i w jakim środowisku chemicznym funkcjonują. Dzięki nowoczesnym metodom obliczeniowym (w tym teorii funkcjonału gęstości, DFT), możemy dziś generować takie widma komputerowo, bez potrzeby wykonywania kosztownych eksperymentów.

Projekt łączy te dane z zaawansowanymi sieciami neuronowymi – **grafowymi sieciami neuronowymi (GNN)**, które przypominają sposób, w jaki neurony w mózgu przetwarzają informacje o powiązaniach, oraz **konwolucyjnymi sieciami neuronowymi (CNN)**, wyspecjalizowanymi w rozpoznawaniu wzorców. Dzięki temu stworzymy hybrydowe modele, które nie tylko przewidują aktywność biologiczną cząsteczek, ale również potrafią ją wyjaśnić.

Modele będą trenowane do rozpoznawania aktywności wobec trzech ważnych receptorów związanych z ośrodkowym układem nerwowym:

- **receptora serotoniny 5-HT₆**, odpowiedzialnego m.in. za pamięć i funkcje poznawcze,
- **receptora μ-opioidowego (MOR)**, kluczowego w regulacji bólu i mechanizmach uzależnienia,
- oraz **receptora histaminowego H₃**, odgrywającego rolę w czuwaniu, śnie i neurodegeneracji.

Są to cele trudne – nie tylko dlatego, że są biologicznie złożone, ale też dlatego, że drobne zmiany w strukturze cząsteczki mogą prowadzić do dramatycznie różnych efektów biologicznych (zjawisko tzw. „klifów aktywności”). Tradycyjne metody często sobie z tym nie radzą – nasze modele, wzbogacone o dane spektroskopowe, mają szansę uchwycić te niuanse.

Co więcej, projekt nie ogranicza się do przewidywania „czy coś działa”. Dzięki metodom wyjaśnialnej AI (np. SHAP), jesteśmy w stanie wskazać, **która część widma lub który fragment cząsteczki** odpowiada za przewidywaną aktywność. To nie tylko zwiększa zaufanie do modelu, ale daje chemikom realne wskazówki do projektowania lepszych związków.

W efekcie projekt ma szansę zrewolucjonizować sposób, w jaki prowadzimy wirtualny screening związków chemicznych – czyniąc go bardziej trafnym, przejrzystym i opartym na rzeczywistej chemii, a nie tylko statystycznej korelacji. To krok w stronę **sztucznej inteligencji, która rozumie chemię**, a nie tylko ją symuluje.