Metody uczenia maszynowego w doborze katalizatorów oraz przewidywaniu reaktywności i enancjoselektywności w katalizie asymetrycznej kompleksami magnezu

mgr Paulina Baczewska

promotor: prof. dr hab. Jacek Młynarski

Tempo, w jakim przybywa danych chemicznych, dawno przekroczyło możliwości ich ręcznego przeglądania. Wiedza zawarta w opublikowanych reakcjach jest rozproszona po setkach artykułów i materiałów uzupełniających, a zależności widoczne dopiero na poziomie większego zbioru mogą łatwo pozostać niezauważone.

Właśnie w tym miejscu widzę rolę uczenia maszynowego. Nie jako wartości samej w sobie, lecz jako narzędzia pozwalającego wydobyć z istniejących wyników zależności użyteczne dla katalizy asymetrycznej, która nadal pozostaje obszarem pełnym wyzwań.

Niniejsza rozprawa doktorska opiera się na założeniu, że o wartości modelu decyduje to, czy pytanie zadane modelowi odpowiada informacji rzeczywiście zawartej w danych i ich reprezentacji. Nie można na przykład pytać o wpływ typu rozpuszczalnika na wynik reakcji, jeżeli w zbiorze nie ma przykładów tej samej reakcji prowadzonej w różnych rozpuszczalnikach. Pytania badawcze stawiałam w obszarze asymetrycznej katalizy kompleksami magnezu.

Pierwszy projekt pokazuje pracę z dużym i niejednorodnym zbiorem danych, którego samo ujednolicenie było istotnym wyzwaniem. Rozproszoną literaturę katalizy magnezowej przekształciłam w spójny zbiór danych, obejmujący również wyniki o niskim nadmiarze enancjomerycznym oraz przykłady niekorzystne, odzyskane ręcznie z materiałów uzupełniających. Zbudowany na tym zbiorze model wskazywał ligandy, które w połączeniu z odpowiednią solą magnezową miały największą szansę dać indukcję asymetryczną dla zadanej reakcji. Jego wartość potwierdziłam eksperymentalnie. Rekomendowane układy poprawiły wyniki znanych redukcji α-ketoestrów, pozwoliły otrzymać produkty, dla których wcześniej nie opisano redukcji katalizowanej kompleksem magnezu, a w części reakcji umożliwiły zastąpienie labilnego odczynnika magnezoorganicznego wygodniejszą w użyciu i tańszą solą magnezu. Praca została opublikowana w *Angewante Chemie*.[1]

Drugi projekt dotyczył sytuacji odwrotnej. Punktem wyjścia były małe, ale jednorodne zbiory danych dla dwóch reakcji opracowanych wcześniej w naszym laboratorium, czyli epoksydacji chalkonów i addycji tio-Michaela. Dane zostały uzupełnione o rzadko publikowane wyniki negatywne. W takim przypadku pytanie ilościowe bywa możliwe, ale nie zawsze. W epoksydacji chalkonów wartości enancjoselektywności były do siebie zbliżone, a brak produktu nie mógł być traktowany jako 0% ee, dlatego sens miało przede wszystkim pytanie o reaktywność. W addycji tio-Michaela, obejmującej szerszą przestrzeń substratów i większe zróżnicowanie wyników, możliwe było przewidywanie nadmiaru enancjomerycznego oraz uchwycenie nieoczywistych trendów. Oba typy predykcji sprawdziłam prospektywnie na substratach niewykorzystanych do budowy modeli. Praca została opublikowana w *ACS Catalysis*.[2]

Obie prace prowadzą do wspólnego wniosku. O użyteczności modelu nie decyduje wyłącznie liczba danych ani złożoność algorytmu, lecz również to, czy postawione pytanie mieści się w granicach informacji zawartej w zbiorze. Tak rozumiany model nie zastępuje eksperymentu. Pomaga wskazać, co warto sprawdzić w pierwszej kolejności i gdzie predykcji należy ufać ostrożnie. Dlatego każdą rekomendację traktowałam jako hipotezę wymagającą weryfikacji doświadczalnej. Uczenie maszynowe

---

[1] Baczewska, P.; Kulczykowski, M.; Zambroń, B.; Jaszczewska-Adamczak, J.; Pakulski, Z.; Roszak, R.; Grzybowski, B. A.; Mlynarski, J. Machine Learning Algorithm Guides Catalyst Choices for Magnesium-Catalyzed Asymmetric Reactions. *Angew. Chemie Int. Ed.* **2024**, *63* (37), e202318487. https://doi.org/https://doi.org/10.1002/anie.202318487.

[2] Baczewska, P.; Nowak, D.; Jaszczewska-Adamczak, J.; Bachorz, R. A.; Hoffmann, M.; Mlynarski, J. Predicting Reactivity and Enantioselectivity from Small Experimental Datasets: A Case Study in Asymmetric Magnesium Catalysis. *ACS Catal.* **2026**. https://doi.org/10.1021/acscatal.6c05090.

może w ten sposób przekształcić opublikowaną już wiedzę w dobrze uzasadnione punkty wyjścia do optymalizacji metod katalizy asymetrycznej.