

ОПТИМИЗАЦИЯ УСЛОВИЙ ЭЛЕКТРОКАТАЛИТИЧЕСКОГО СИНТЕЗА НА ОСНОВЕ АЛГОРИТМОВ МАШИННОГО ОБУЧЕНИЯ

Рудакова М.Д. (ИТМО), Алиев Т.А. (ИТМО), Семенов А.В.

Научный руководитель – кандидат технических наук, Мешков А.В.
(ИТМО)

Введение. Электроорганический синтез на основе переменного тока демонстрирует повышенную эффективность проведения окислительно-восстановительных реакций [1]. Использование переменного тока совместно с никелевыми катализаторами позволяет существенно упростить и снизить стоимость синтеза молекул с гетероатомными связями, производные которых используются в качестве лекарственных средств [2].

На выход электрокаталитических реакций влияет множество факторов, прежде всего природа реагирующих веществ, катализатора, растворителя, электролита и т. д., а также параметры и тип тока, концентрация веществ, температура и время синтеза.

Традиционный подход к оптимизации условий синтеза экспериментальным путем занимает значительное время. Интеграция алгоритмов машинного обучения может существенно ускорить этот процесс, став при этом удобным инструментом при планировании экспериментов [3].

Основная часть. Экспериментальные данные электрокатализа содержат большое количество названий молекул, химия которых играет важную, иногда даже определяющую роль в протекании реакции. Для того, чтобы заменить категориальные значения, при этом численно описать химические свойства и структуру молекул были использованы квантово-химические величины, рассчитанные по теории функционала плотности (англ. density functional theory, DFT): E_{HOMO} (англ. highest occupied molecular orbital) – энергия высшей заполненной молекулярной орбитали; E_{LUMO} (англ. lowest unoccupied molecular orbital) – энергия низшей свободной молекулярной орбитали, характеризует электрофильные свойства молекулы; дипольный момент, характеризует электрические свойства. В итоговом датасете все названия молекул были заменены на расчётные данные DFT: энергию HOMO, энергию LUMO и дипольный момент.

Для прогнозирования оптимальных параметров синтеза были построены регрессионные модели, базирующихся на библиотеке Scikit-learn: простая линейная регрессия, а также ансамблевые методы Random Forest и Gradient Boosting. Для увеличения предсказательной эффективности была произведена оптимизация гиперпараметров методом GridSearch.

Выводы. В результате была получена модель машинного обучения, которая позволяет с высокой эффективностью прогнозировать результаты электрохимических реакций. Это можно использовать непосредственно для оценки ожидаемого выхода перед началом эксперимента, а также для оптимизации условий реакции, при которых выход будет наибольшим.

Список использованных источников:

1. Zeng L., Wang J., Wang D. et al. Comprehensive Comparisons between Directing and Alternating Current Electrolysis in Organic Synthesis // *Angewandte Chemie International Edition* – 2023.
2. Evgeniy O. Bortnikov and Sergey N. Semenov Coupling of Alternating Current to Transition-Metal Catalysis: Examples of Nickel-Catalyzed Cross-Coupling // *The Journal of Organic Chemistry*. – 2021. – V. 86. - №1. – P. 782-793.
3. Steinmann S. N., Wang Q., She Z. W. How machine learning can accelerate electrocatalysis discovery and optimization // *Materials Horizons*. – 2023. – V.10, – P. 393-406