

Тема работы: разработка метода для определения распределений пор по размерам функциональных материалов с помощью нейронных сетей, обученных на комбинированном наборе данных.

Состав коллектива:

- Иванов Антон Дмитриевич, выпускник магистратуры факультета информационных технологий НГУ, лаборант ИК СО РАН, a.ivanov10@g.nsu.ru
- Гренев Иван Васильевич, к. ф.-м. н, в.н.с. ИК СО РАН, greneviv@gmail.com

Финансовая поддержка: Грант РФФИ 24-71-10096 “Новый многоуровневый подход к разработке адсорбционных материалов для процессов газоразделения на основе методов молекулярного моделирования и машинного обучения”, рук. к. ф.-м. н. Гренев Иван Васильевич.

Научное содержание работы:

Постановка задачи: Целью работы является разработка и валидация нейросетевого подхода, предназначенного для определения дифференциальной функции распределения пор по размерам на основе экспериментальных данных физической адсорбции азота. Для достижения поставленной цели необходимо было произвести массовую предобработку накопленных экспериментальных данных, генерацию репрезентативного синтетического обучающего датасета из сотен тысяч изотерм адсорбции, сконструировать и обучить набор нейросетевых моделей различных архитектур.

Современное состояние проблемы: на момент начала проведения работы общепринятым стандартом решения данной задачи являлось использование классических математических методов, основанных на минимизации функционала невязки интегрального уравнения Фредгольма первого рода с применением регуляризации Тихонова. Главный недостаток этих подходов заключается в высокой степени субъективности при выборе вида сглаживающего функционала и веса регуляризации, что нередко приводит к искусственному размытию реальных узких пиков распределений или появлению ложных физически необоснованных осцилляций. В качестве перспективной

альтернативы выступает идея применения искусственных нейронных сетей. Несмотря на то, что в настоящее время нейросети активно используются для решения обратных задач в различных смежных областях, в литературе отсутствовали данные о систематическом применении данного подхода к задаче определения распределений пор по размерам.

Подробное описание работы, включая используемые алгоритмы: Весь вычислительный процесс в рамках исследования был разбит на несколько этапов и реализован в рамках программной экосистемы языка программирования Python. На первом этапе была проведена масштабная автоматизированная конвертация, фильтрация и интерполяция базы данных, включающей в себя 26000 реальных экспериментальных изотерм адсорбции. Далее с целью последующего сравнения эффективности моделей был реализован классический метод решения обратной задачи определения распределений пор по размерам. Математически данный подход сводится к минимизации функционала невязки дискретного интегрального уравнения Фредгольма первого рода с ограничением на неотрицательность искомой функции распределения пор (использовался алгоритм NNLS из пакета SciPy). Поскольку обратная задача относится к классу математически некорректных задач, в целевую функцию вводился регуляризирующий функционал второго порядка, штрафующий решение за избыточную кривизну. Для автоматического поиска компромиссного параметра регуляризации был разработан алгоритм на основе метода L-кривой, строящий логарифмическую зависимость гладкости решения от нормы невязки. Ввиду того, что на практике график часто имел слабо выраженный перегиб, наиболее устойчивые результаты показала модификация алгоритма, осуществляющая поиск точки, максимально удаленной от хорды, соединяющей концы кривой. Применение этого классического алгоритма к тысячам кривых из базы данных оказалось крайне трудоемким с вычислительной точки зрения процессом, поскольку для каждой изотермы требовалось проводить многократный итерационный поиск оптимального параметра регуляризации, использование вычислительных возможностей кластера помогло решить данную задачу за разумное время.

Для формирования обучающей выборки был разработан алгоритм случайной генерации мультимодальных распределений пор, моделируемых в виде комбинаций нормальных и логнормальных функций.

Полученный массив данных использовался для обучения трех различных архитектур нейронных сетей с использованием фреймворка PyTorch. В ходе работы были сконструированы и обучены: многослойный перцептрон, одномерная сверточная нейросеть, а также гибридная модель на основе сверточного автоэнкодера.

Полученные результаты: в ходе данной работы решена актуальная научно-практическая задача разработки и исследования подходов применения нейросетевых моделей для определения дифференциальных распределений пор по размерам функциональных пористых материалов на основе экспериментальных изотерм адсорбции азота. Был предложен подход, позволяющие обойти ключевые фундаментальные ограничения классических регуляризационных методов оптимизации решения обратной некорректной задачи интеграла адсорбции за счет сужения пространства поиска решений в рамках физически обоснованных закономерностей. Экспериментально показано, что структура синтетической обучающей выборки выступает в роли нелинейного регуляризатора. Нейросетевые модели, обученные на гладких нормальных пиках, успешно подавляют численные осцилляции, присущие классическому решению. Гибридная модель на основе автоэнкодера обеспечила наиболее сбалансированные результаты. За счет учета экспериментальных особенностей при обучении, модель показала высокую точность на реальных объектах ($MRE = 7.86\%$) при сохранении низкого уровня ошибки на синтетической тестовой выборке ($MAE = 0.35$).

Эффект от использования кластера в достижении целей работы: Применение кластера сыграло ключевую роль в успешном выполнении работы. Классический подход, требующий многократного итерационного поиска параметра регуляризации по методу L-кривой для каждой изотермы, оказался крайне трудоемким, распараллеливание расчетов на узлах кластера позволило решить эту задачу за разумное время.

Перечень публикаций, содержащих результаты работы:

1) Иванов, А. Д. Определение распределения пор по размерам адсорбентов и катализаторов с помощью методов машинного обучения / А. Д. Иванов, И. В. Гренев. // Тезисы XXV Всероссийской конференции молодых учёных по математическому моделированию и информационным технологиям (Новосибирск, 21–25 октября 2024 г.). — Новосибирск : ФИЦ ИВТ, 2024. — С. 37. — ISBN 978-5-905569-27-2.

2) Иванов, А. Д. Применение методов машинного обучения для определения распределений пор по размерам адсорбентов и катализаторов / А. Д. Иванов, И. В. Гренев. // Физико-химические проблемы адсорбции, структуры и химии поверхности нанопористых материалов: сборник трудов XXI Всероссийского симпозиума с международным участием, посвященного 125-летию Л. В. Радushкевича (Москва, 22–25 апреля 2025 г.). – Москва: ИФХЭ РАН, 2025. – С. 34–35. – ISBN 978-5-00202-834-4.

3) Иванов, А. Д. Определение распределения пор по размерам из изотерм адсорбции с помощью методов машинного обучения / А. Д. Иванов, И. В. Гренев. // СХС-2025: материалы II Международного Сибирского химического симпозиума (Томск, 20–24 октября 2025 г.). – Томск: Изд-во ТПУ, 2025. – С. 213–214.

4) Machine Learning-Based Predicting of the Equilibrium Cation Distribution in Faujasite-Type Zeolites / I. V. Grenev, M. E. Bobkov, A. D. Ivanov [et al.]. – Текст : электронный // Materials Chemistry and Physics. – 2026. – Vol. 349, Part 2. – Art. 131853, DOI: 10.1016/j.matchemphys.2025.131853.