06

Определение долей атомов платины в агломератах биметаллических наночастиц с помощью методов машинного обучения

© Я.Н. Гладченко-Джевелекис, Д.Б. Толчина, С.В. Беленов, В.В. Срабионян, В.А. Дурыманов, И.А. Викленко, Л.А. Авакян, А.А. Алексеенко, Л.А. Бугаев

Южный федеральный университет,
344090 Ростов-на-Дону, России
e-mail: ygl@sfedu.ru; vvsrab@sfedu.ru

Поступило в Редакцию 24 октября 2024 г.
В окончательной редакции 11 марта 2025 г.
Принято к публикации 1 апреля 2025 г.

Рассмотрена применимость методов машинного обучения, в частности, искусственных нейронных сетей, для получения информации о распределении атомов целевого вещества в составе агрегатов биметаллических наночастиц различных архитектур на основе данных о парных радиальных функциях распределения атомов. Прямыми источниками данных являются экспериментальные методы рентгеновской дифракции и рентгеновской абсорбционной спектроскопии из протяженной энергетической области спектра. Обученная модель искусственной нейронной сети продемонстрировала высокую точность определения долей атомов платины в составе наночастиц различных архитектур в агломерате (коэффициент детерминации ~ 0.98). Апробация обученной модели на экспериментальных данных для катализаторов, содержащих биметаллические наночастицы PtCu, показала ее высокую обобщающую способность, что свидетельствует о перспективности применения такого подхода к определению эффективности расхода платины при синтезе платиносодержащих катализаторов на основе наночастиц.

Ключевые слова: наночастицы типа ядро-оболочка, градиентные наночастицы, парные радиальные функции распределения атомов, спектроскопии из протяженной энергетической области спектра, искусственные нейронные сети, CatBoost.

DOI: 10.61011/JTF.2025.08.60901.365-24

Введение

В настоящее время платиновые катализаторы применяются во многих отраслях производства. В частности, платиносодержащие наночастицы (НЧ), нанесенные на высокодисперсный углеродный материал, активно используются в разработке топливных элементов в качестве катализаторов [1]. Ввиду высокой стоимости платины большое внимание уделяется разработке высокоэффективных и стабильных платиносодержащих катализаторов с более низкой себестоимостью [2–6]. Одним из способов решения этой задачи является использование биметаллических платиносодержащих НЧ с добавлением d-металлов, например никеля, кобальта или меди [1,7]. При этом каталитические свойства таких НЧ зависят как от состава, так и от их архитектуры, т. е. пространственного распределения компонентов по объему НЧ.

Известно, что биметаллические НЧ с архитектурой ядро-оболочка с платиновой оболочкой демонстрируют каталитическую активность, сравнимую с монометаллическими платиновыми НЧ, но при этом содержат значительно меньше дорогостоящей платины [2–4]. Однако следует отметить, что в процессе эксплуатации архитектура таких НЧ может со временем изменяться, например, граница перехода от ядра к оболочке может становиться менее выраженной в результате термического воздействия [8,9], что при условии тонкого слоя оболочки может приводить к ухудшению каталитических свойств НЧ. Недавние исследования показали, что наличие плавного переходного слоя между областями ядра и оболочки позволяет добиться улучшенной стабильности катализатора [10,11]. Такую архитектуру НЧ предложили называть „градиентной“ [10,12,13].

Несмотря на всю перспективность „градиентной“ архитектуры НЧ, в процессе многостадийного синтеза допускается формирование НЧ других типов, будь то моно- или биметаллические НЧ различных архитектур. Поэтому для производства НЧ заданной архитектуры с определенными свойствами необходима существенная оптимизация условий синтеза, которая требует наличия структурной информации о НЧ. Наиболее распространенными источниками получения такой информации являются спектроскопия рентгеновского поглощения (XAS), методы рентгеновской дифракции (XRD) и электронной микроскопии высокого разрешения с элементным картированием, а также некоторые другие методы [14,15].

Применение методов машинного обучения способно существенно упростить и ускорить извлечение и анализ структурной информации о нанообъектах. Например, было продемонстрировано, что использование нейронных сетей позволяет извлекать информацию о координационных числах и расстояниях из XAS данных [16,17], а

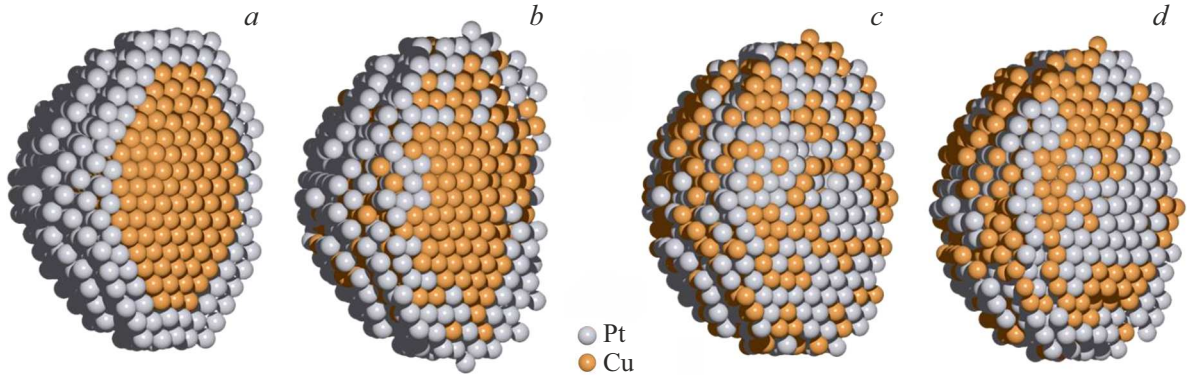

Рис. 1. Рассматриваемые архитектуры биметаллических НЧ: *a* — ядро-оболочка, *b* — „градиент“, *c* — неупорядоченный и *d* — агрегированный сплавы.

также позволяет с высокой точностью определять архитектуру НЧ из данных о парных радиальных функциях распределения атомов (ПРФРА) [18].

Настоящая работа посвящена исследованию принципиальной возможности определения эффективности расхода платины во время синтеза, т. е. определение доли платины, расходуемой на образование НЧ с конкретной архитектурой.

1. Методы и подходы

1.1. Источники получения ПРФРА

Основными способами получения ПРФРА является анализ данных рентгеновской дифракции и рентгеновской абсорбционной спектроскопии из протяженной энергетической области спектра (EXAFS). В основе метода извлечения ПРФРА из XRD лежит зависимость функции радиального распределения атомной плотности от интенсивности когерентного рассеяния рентгеновских лучей при дифракции [19]. Из-за физических ограничений измерения рассеяния рентгеновских лучей, при извлечении радиальных функций распределения из XRD возникают краевые эффекты, проявляющиеся в появлении ложных максимумов в кривых функции распределения. Получение сведений о ПРФРА из данных EXAFS основано на многопараметрической оптимизации фурье-образа F(R) исходного сигнала. Такой анализ позволяет для средней по образцу, репрезентативной НЧ с высокой точностью определять радиусы координационных сфер (~ 0.01 Å) [20], и с существенно меньшей точностью координационные числа (погрешность $\sim 10\,\%$), а также параметр температурного и структурного разупорядочения [21,22]. Следует отдельно отметить, что процесс EXAFS-анализа может существенно различаться в зависимости от исследуемого материала, а также, что надежность определения структурных параметров при рассмотрении более удаленных координационных сфер уменьшается. В настоящей работе используются ПРФРА, рассчитанные напрямую для молекулярных нанокластеров, поэтому при сопоставлении с экспериментальными данными для корректности результатов необходима дополнительная нормировка координационных чисел, определенных из EXAFS-анализа, что обусловлено тем фактом, что атомы металла могут находиться в двух состояниях, принадлежащих двум компонентам: наночастицам и оксиду:

$$N_{A-B}^{Np} = \frac{V}{V - N_{A-O}^{\mathrm{EXAFS}}} \cdot N_{A-B}^{\mathrm{EXAFS}}, \qquad (1)$$

где N_{A-B}^{EXAFS} — координационное число атомов сорта B относительно сорта A, полученное из EXAFS-анализа; V — предполагаемое координационное число для атомов металла сорта A в окисленном состоянии, которые могут находиться в материале не в составе НЧ (в настоящей работе предполагается, что $V = 6$); N_{A-O}^{EXAFS} — координационное число атомов кислорода относительно атомов сорта A. Более подробный вывод выражения (1) приведен в Приложении 1.

1.2. Методы машинного обучения. Введение в задачу, используемые метрики

Поставленная задача определения эффективности расхода платины при синтезе НЧ заданной архитектуры может быть сведена к определению долей платины, ушедших на формирование НЧ различных архитектур в их агломерате. Вероятность образования конкретной архитектуры НЧ существенно зависит от процедуры синтеза [6,9–11,23], поэтому мы ограничились рассмотрением наиболее вероятных архитектур НЧ, образующихся в процессе синтеза платиносодержащих катализаторов. Таким образом, в настоящей работе рассматривались монометаллические НЧ платины и биметаллические НЧ состава PtCu со следующими архитектурами: медное ядро и платиновая оболочка, платина-медная НЧ со структурой сплава и агрегированного сплава, и градиентная НЧ с медным ядром и платиновой оболочкой.

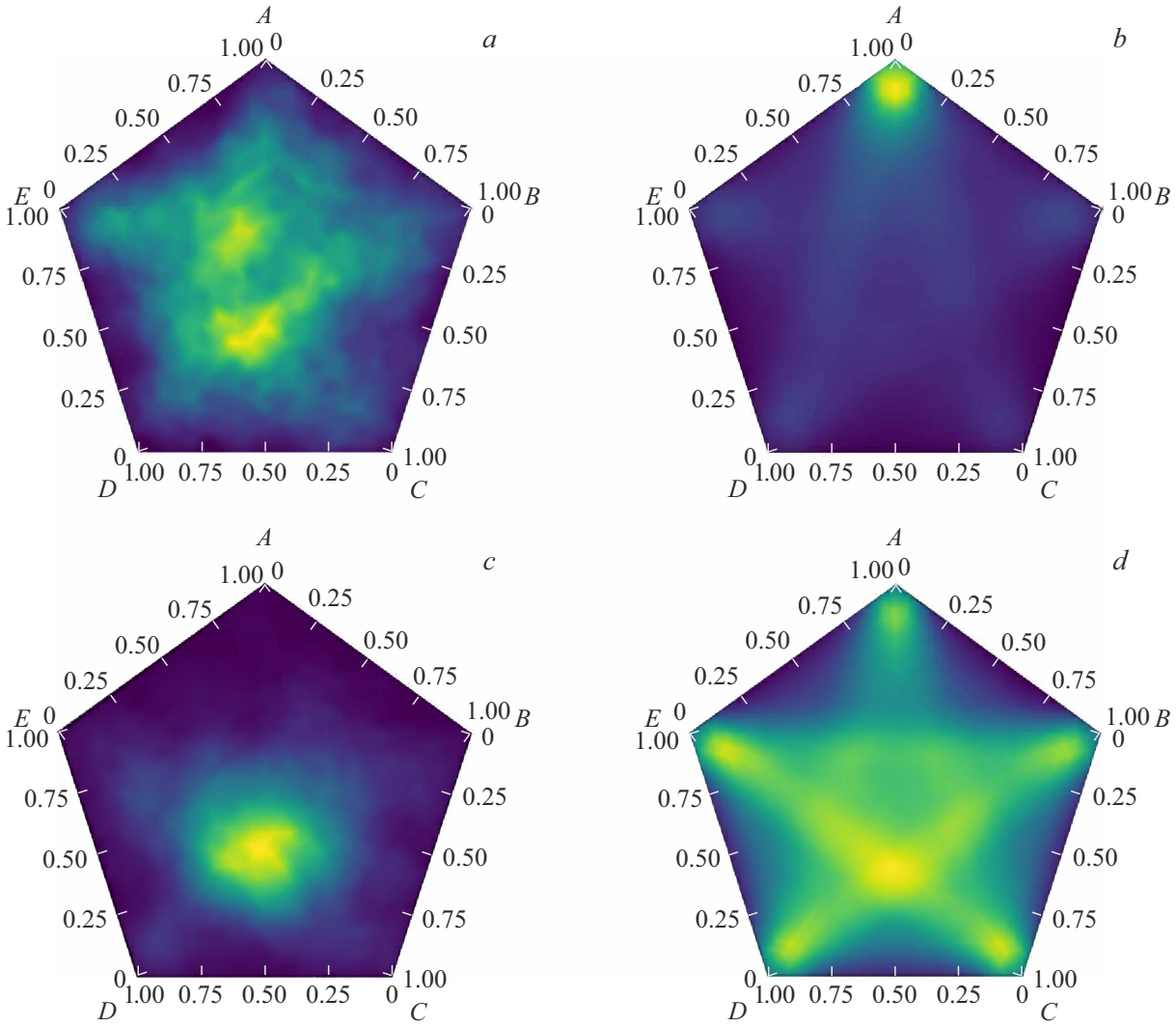

Рис. 2. a, c — распределение долей НЧ различных архитектур; b, d — распределение долей атомов платины в ансамбле из пяти НЧ. A — Pt, B — PtCu, C — Cu@Pt, D — PtCu_{aggr}, E — CuPt_{grad}.

Изображения рассматриваемых архитектур биметаллических НЧ приведены на рис. 1.

Таким образом, поставленная задача сводится к задаче многоцелевой регрессии на табличных данных, которая может быть решена методами машинного обучения с учителем. Формально работа модели может быть представлена следующим образом:

$$\Phi(x^i, \Theta) \to \left\{c_{\mathrm{Pt}}, c_{\mathrm{PtCu}}, c_{\mathrm{Cu@Pt}}, c_{\mathrm{PtCu_{aggr}}}, c_{\mathrm{CuPt_{grad}}}\right\}^i,$$

где x^i — ПРФРА для i-го экземпляра данных, Θ — массив параметров модели, $\{c_{\mathrm{Pt}}, c_{\mathrm{PtCu}}, c_{\mathrm{Cu@Pt}}, c_{\mathrm{PtCu_{aggr}}}, c_{\mathrm{CuPt_{grad}}}\}$ — предсказанный моделью вектор долей платины в составе НЧ рассматриваемых архитектур в агломерате. В процессе обучения модели подбираются значения массива параметров Θ таким образом, чтобы приблизить предсказанные моделью и эталонные значения долей атомов платины, согласно некоторой функции потерь. В настоящей работе в качестве функции потерь использовалась функция среднеквадратичного отклонения (MSE), определяемая выражением:

$$MSE = \frac{1}{n}\Sigma_{i=1}^{n}(y_i - y_i^*)^2,$$

где y_i и y_i^* — истинный и предсказанный моделью вектор целевых значений соответственно.

2. Результаты и обсуждения

2.1. Формирование обучающего набора данных

В настоящей работе рассматриваются ПРФРА платиновых монометаллических и платина-медных биметаллических НЧ сферической формы с четырьмя различными архитектурами (рис. 1). Атомные модели НЧ полу-

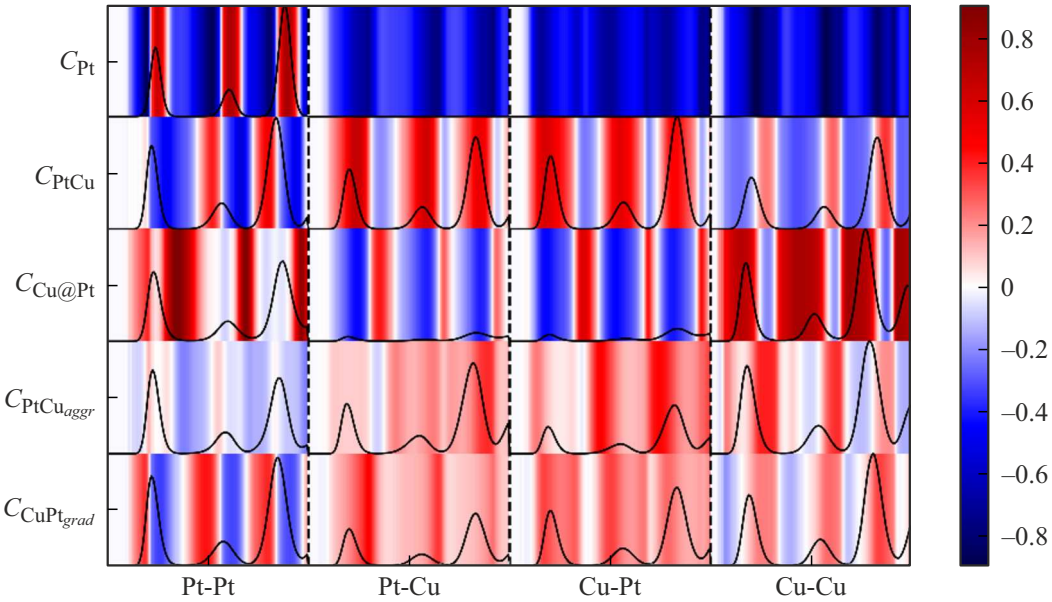

Рис. 3. Матрица корреляций между входными и выходными параметрами. Для каждой архитектуры приведен характерный набор парных радиальных функций распределения атомов, являющийся входными данными.

чались путем вырезания сферической области из бесконечного ГЦК кристалла с последующим молекулярно-динамическим моделированием, как это было описано в работе [18]. Всего было получено 1456 различных платина-медных биметаллических НЧ с шестнадцатью различными размерами (от 1.3 до 6 nm) и тринадцатью соотношениями компонент Pt:Cu (от 20 : 80 до 80 : 20) и 22 платиновые монометаллические НЧ размерами от 1.4 до 9.5 nm.

Ансамбль из пяти НЧ различных архитектур характеризуется ПРФРА (PRDF), полученной по формуле (2), где c_{arch} — доля вклада НЧ данной архитектуры в ансамбле, n^{Pt}_{arch} — количество атомов платины в наночастице данной архитектуры, PRDF_{arch} — ПРФРА конкретной НЧ заданной архитектуры

$$\mathrm{PRDF} = \frac{\Sigma_{arch} c_{arch} \cdot \mathrm{PRDF}_{arch} \cdot n^{\mathrm{Pt}}_{arch}}{\Sigma_{arch} c_{arch} \cdot n^{\mathrm{Pt}}_{arch}}. \qquad (2)$$

Для формирования набора данных такого типа необходимо случайным образом сгенерировать набор концентраций carch, сумма которых будет равна единице в рамках одного ансамбля. Данные коэффициенты были сгенерированы с помощью функции Дирихле с параметрами $\alpha_1 = [0.6, 0.6, 0.6, 0.6, 0.6]$. Полученное в результате такой операции распределение carch представлено на рис. 2, *a*. Полученный данные имеют большую плотность в окрестности точки *A* — доли платины в составе монометаллической платиновой НЧ (рис. 2, *b*). Это связано с тем, что у нас ограниченный набор монометаллических НЧ с высоким средним значением количества атомов платины в НЧ. При суммировании ПРФРА (формула (2)) наночастиц различных архитектур мы с большей вероятностью выбираем крупные монометаллические платиновые НЧ, соответственно, доля платины, „расходующаяся“ на их образование, больше, чем все остальные. Чтобы исправить подобное распределение, были объединены два набора данных, полученных в результате генерации коэффициентов c_{arch} с помощью функций Дирихле с параметрами $\alpha_2 = [0.2, 0.6, 0.6, 0.6, 0.6]$ и $\alpha_3 = [0.5, 3, 3, 3, 3]$ (рис. 2, *c*).

Для каждого набора из пяти концентраций carch случайно выбираются по 208 НЧ каждой архитектуры с различными размерами и соотношениями компонент Pt : Cu, и ПРФРА ансамбля вычисляется по формуле (2). Таким образом формируется обучающая выборка, суммарно состоящая из 208 000 строк.

Интересующие нас сведенья о долях атомов платины, расходуемых на образование НЧ определенной архитектуры в ансамбле c^{Pt}_{arch}, определяется, как

$$c^{\mathrm{Pt}}_{arch} = \frac{c_{arch} \cdot n^{\mathrm{Pt}}_{arch}}{\Sigma_{arch} c_{arch} \cdot n^{\mathrm{Pt}}_{arch}}. \qquad (3)$$

Полученное в результате распределение c^{Pt}_{arch} представлено на рис. 2, *d*. Результирующий набор обучающих данных был разделен на 3 части: тренировочный набор — 60 % от всей выборки; валидационный набор — 15 % от выборки, и тестовый набор — 25 % выборки. Согласно общепринятой парадигмы, тренировочный набор данных использовался для обучения матриц параметров Θ используемых моделей машинного обучения, валидационная выборка использовалась для подбора гиперпараметров моделей и определения момента остановки обучения, а тестовая выборка использовалась для несмещенной оценки качества работы модели.

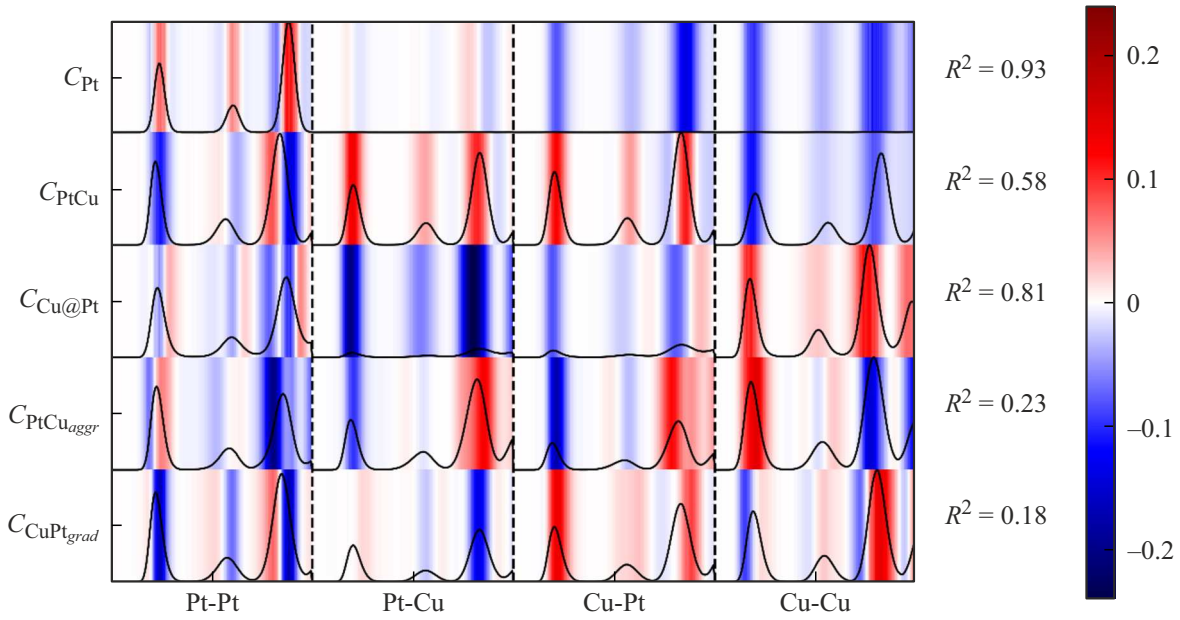

Рис. 4. Матрица коэффициентов модели персептрона. Для каждой архитектуры приведен характерный набор парных радиальных функций распределения атомов, являющийся входными данными.

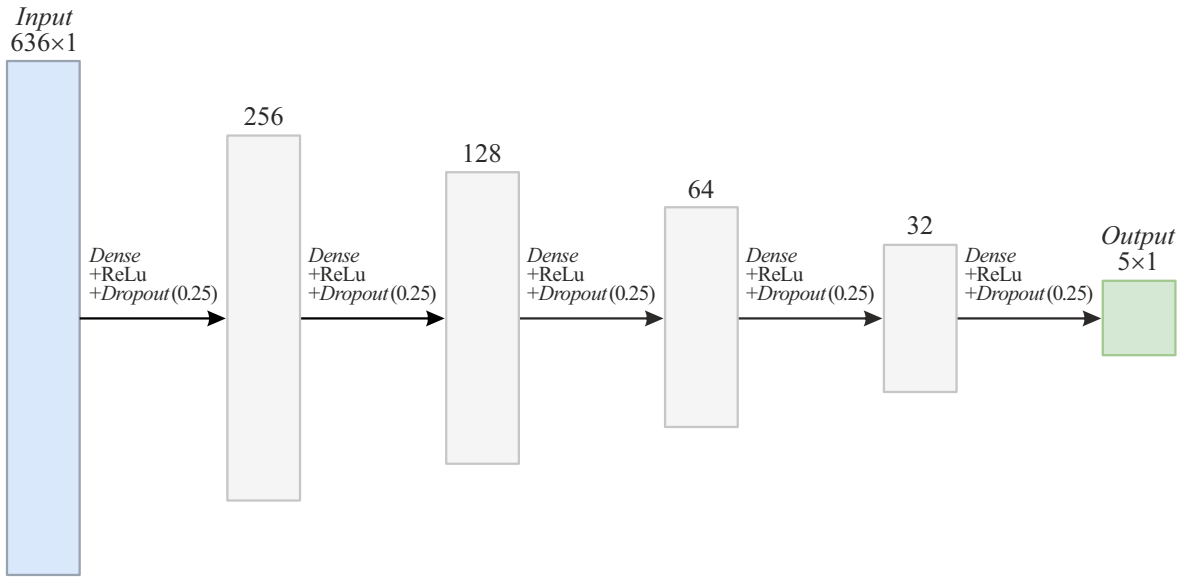

Рис. 5. Схема нейронной сети.

2.2. Выбор моделей

При построении матрицы корреляций входных данных было обнаружено, что между многими признаками существуют значительные корреляции. Эти корреляции обусловлены структурой входной функции, представляющей собой гауссовы пики. Имеющая место мультиколлинеарность данных накладывает ограничения на выбор возможных моделей или на предобработку данных. На рис. 3 представлена матрица корреляций между входными и выходными параметрами.

С учетом особенностей входных данных в качестве базовой модели машинного обучения использовалась модель гребневой регрессии (ridge regression). Для нее были достигнуты высокие средние значения коэффициента детерминации, составившие в среднем ~ 0.98. Однако предсказанные доли атомов платины в агломератах биметаллических наночастиц при проверке модели на экспериментальных данных не имеют физического смысла. Кроме того, распределение остатков содержало значительные выбросы. В связи с этим была рассмотрена модель персептрона, реализующая линейную регрессию

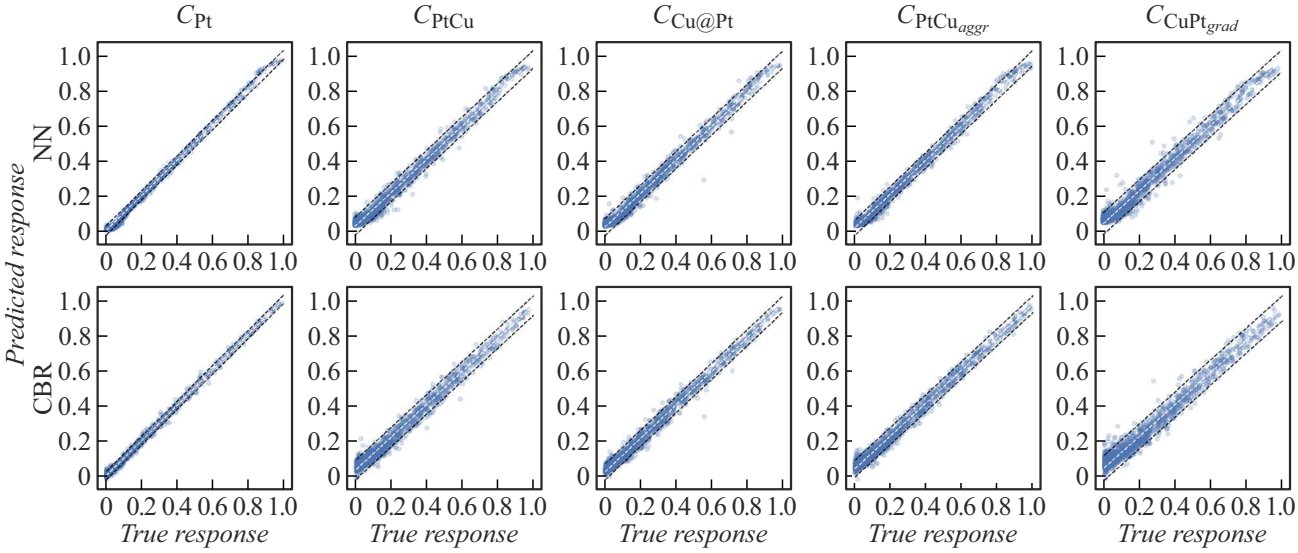

Рис. 6. Сопоставление истинных и предсказанных значений долей атомов платины в ансамбле из пяти НЧ для ИНС (NN) и CBR.

с дополнительным применением сигмоидной функции к выходным данным. Это позволило избежать отрицательных значений при предсказании целевых параметров, но при этом значения коэффициента детерминации упали до 0.58, 0.23 и 0.12 для моделей, определяющих долю атомов платины в составе неупорядоченных сплавов, агрегированных сплавов и „градиентных“ наночастиц соответственно. На рис. 4 представлена матрица коэффициентов модели персептрона с видом типичных функций парного радиального распределения атомов для данных архитектур наночастиц.

Известно, что модели на основе градиентного бустинга на решающих деревьях могут демонстрировать как высокую точность, так и высокую обобщающую способность при работе с табличными данными [24], поэтому в настоящей работе мы ограничились рассмотрением методов градиентного бустинга и искусственных нейронных сетей (ИНС) при решении поставленной задачи.

В качестве модели градиентного бустинга на решающих деревьях использовался CatBoostRegressor (CBR), реализованный в библиотеке CatBoost [25,26], а в качестве альтернативного подхода использовалась полносвязная нейронная сеть с архитектурой, представленной на рис. 5. Подробности о настройке гиперпараметров используемых моделей приведены в Приложении 2.

Описанная модель нейронной сети принимает на вход вектор четырех последовательно соединенных ПРФРА: Pt-Pt, Pt-Cu, Cu-Pt, Cu-Cu, после чего данные последовательно проходят через полносвязные слои с функцией активации ReLU и слоем регуляризации Dropout, который зануляет 25% выходных значений. На последнем слое после функции активации ReLU применяется функция Softmax, которая перенормирует выходной вектор таким образом, чтобы сумма его компонент равнялась единице.

Модель на основе градиентного бустинга состояла из деревьев решения глубиной 10 и количеством шагов бустинга, равным 1000, коэффициент L2 регуляризации равен 3, функция потерь — MultiRMSE [26].

Известно, что некоторые методы машинного обучения могут быть чувствительны к масштабу входных данных или наличию корреляции. Поэтому при тренировке моделей рассматривались дополнительная стандартизация входных данных для их приведения к единому масштабу и применение метода главных компонент (PCA) для сокращения размерности входных данных и получения некоррелированных входных признаков. Однако такие модификации входных данных не привели к сколь-либо заметному улучшению качества работы ИНС или модели градиентного бустинга относительно поставленной задачи.

3. Результаты обучения

3.1. Синтетические данные

Результаты работы обученных моделей на синтетических тестовых данных приведены на рис. 6, а статистика распределения ошибок — на рис. 7. Сравнение распределения остатков для моделей ИНС, обученных на наборе данных, соответствующем параметрам генерации α_1, и расширенном наборе данных приведено на рис. 8.

Несмотря на то что обе модели демонстрируют практически одинаковое качество описания синтетических данных, при сопоставлении с экспериментальными данными ИНС показывает систематически более высокую обобщающую способность.

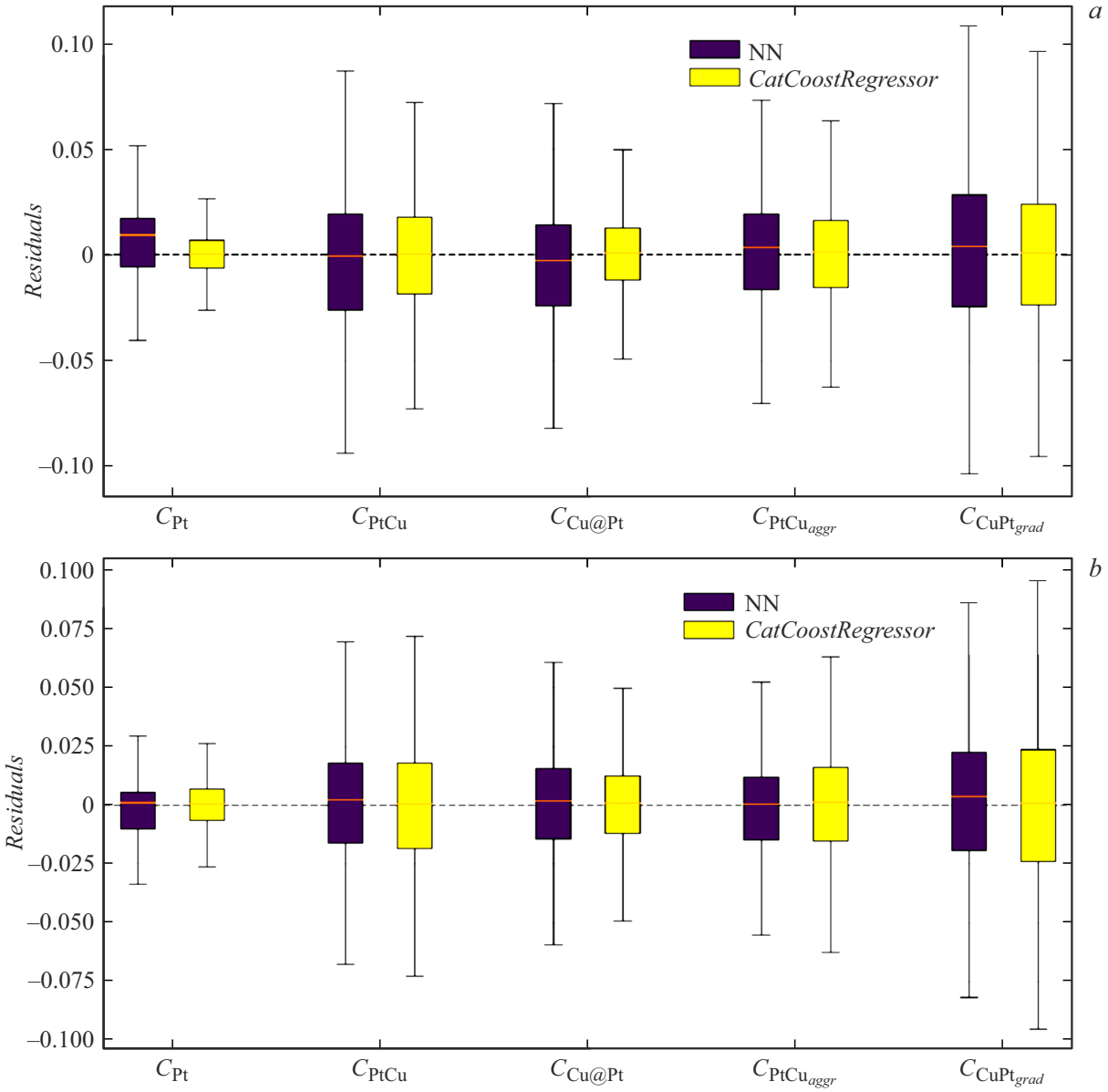

Рис. 7. Сравнение распределения остатков для ИНС (NN) и CBR: *a* — модели, обученные на наборе данных, соответствующем параметрам генерации α_1; *b* — модели, обученные на расширенном наборе данных.

3.2. Сопоставление с экспериментальными данными

Для проверки применимости обученных моделей на основе ИНС и CBR к реальным данным были дополнительно рассмотрены результаты из работы [10], в которой авторы использовали процедуру многостадийного синтеза для получения „градиентных“ платина-медных НЧ. Схематичное представление процедуры синтеза приведено на рис. 9, заимствованном из статьи [10].

Материалы, полученные на 2-й, 3-й и 4-й стадиях, обозначаются как „PtCu_stage2“, „PtCu_stage3“ и „PtCu_stage4“ соответственно. Также в работе С.В. Беленова и др. [11] дополнительно были проведены исследования как одновременного осаждения атомов платины и меди с ожидаемым образованием НЧ с архитектурой твердого раствора, обозначаемые как „PtCu_sim“, так и последовательного двухэтапного осаждения меди и платины с ожидаемым образованием частиц с медным ядром и платиновой оболочкой, которые обозначались как „PtCu_seq“.

Для корректной работы обученных моделей, данные ПРФРА, полученные из EXAFS анализа в работах [10,11], были нормированы в соответствии с формулой (1). Результаты работы моделей представлены на рис. 10.

Из приведенных результатов видно, что модель CBR указывает на наличие существенной доли атомов Pt в составе монометаллических НЧ во всех исследуемых

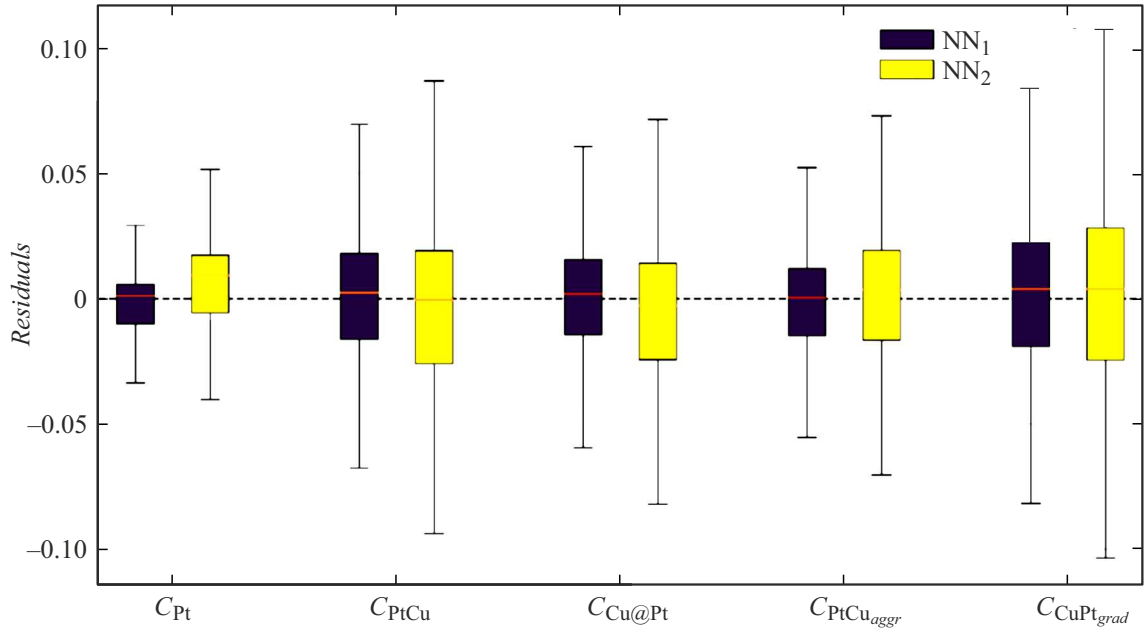

Рис. 8. Сравнение распределения остатков для NN_1 (модель, обученная на расширенном наборе данных) и NN_2 (модель, обученная на наборе данных, соответствующем параметрам генерации α_1).

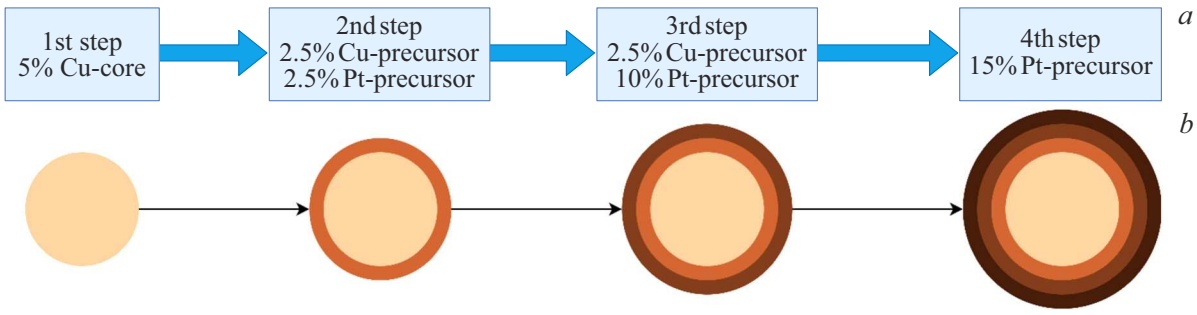

Рис. 9. *a* — схема поэтапного синтеза материалов PtCu/C; *b* — схематическое изображение архитектуры НЧ PtCu, образующихся на каждом из четырех этапов последовательного синтеза градиентного катализатора PtCu/C.

образцах. С другой стороны, модель ИНС свидетельствует о незначительной доли монометаллических НЧ Pt в образцах PtCu_sim и PtCu_seq. Для образца PtCu_sim модель CBR указывает на примерно равномерное содержание атомов Pt в составе рассматриваемых архитектур, за исключением архитектуры неупорядоченного сплава, для которого модель предсказывает меньшее содержание Pt. Архитектура неупорядоченного сплава действительно может оказаться нестабильной, поскольку ГЦК структуры объемных образцов Pt и Cu характеризуются заметно различными параметрами решетки, что делает архитектуру агрегированного сплава более вероятной по отношению к архитектуре неупорядоченного раствора. Поэтому ожидаемо, что при последовательном осаждении компонентов в образце PtCu_seq вероятность образования НЧ с архитектурой неупорядоченного сплава будет не больше, чем при одновременном осаждении компонентов в образце PtCu_sim. Однако, согласно результатам использования модели CBR, доля атомов Pt в составе НЧ архитектуры неупорядоченного сплава (32 %) максимальна и существенно превышает аналогичное значение для PtCu_sim. Одновременно с этим модель CBR для образца PtCu_seq указывает на относительную малость доли атомов Pt в составе НЧ с архитектурами „ядро-оболочка“ и „градиент“, что с учетом процедуры синтеза и обсуждаемых ранее результатов свидетельствует о недостаточной обобщающей способности модели CBR по отношению к экспериментальным данным.

В то же время для образца PtCu_sim модель ИНС указывает на то, что большая часть атомов платины ($\sim$ 88 %) расходуется на образование НЧ с архитектурой агрегированного сплава. В образце PtCu_seq большая часть платины находится в составе НЧ с архитектурой ядро-оболочка ($\sim$ 66 %), и значительная часть ($\sim$ 30 %) — в составе НЧ с архитектурой агре-

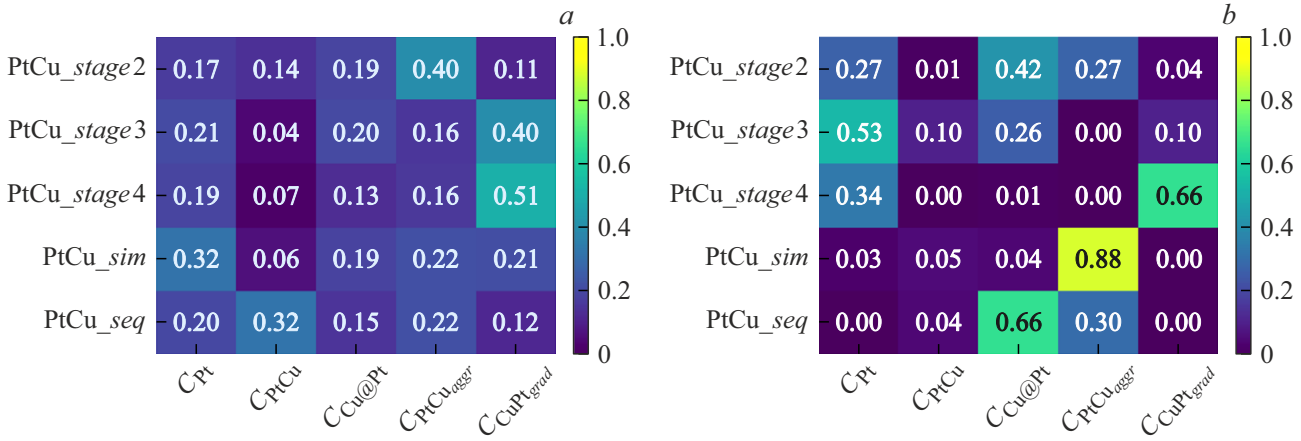

Рис. 10. Результат использования моделей CBR (*a*) и ИНС (*b*) на экспериментальных данных.

гированного сплава. Для образцов, полученных на 2-й, 3-й и 4-й стадиях синтеза, модель ИНС указывает на монотонное увеличение доли атомов Pt в составе НЧ „градиентной“ архитектуры, с одновременным монотонным уменьшением НЧ с архитектурой „ядро-оболочка“. При этом на этих стадиях синтеза наблюдается существенный вклад монометаллических платиновых НЧ. Наблюдаемые результаты применения модели ИНС для всех рассматриваемых образцов представляются ожидаемыми и логичными, и находятся в полном согласии с результатами работы [10,11].

Заключение

В работе продемонстрирована принципиальная возможность определения долей атомов целевого вещества в составе агрегатов наночастиц различных архитектур с применением методов машинного обучения и, в частности, искусственных нейронных сетей по данным парных радиальных функций распределения атомов. Обученная модель ИНС демонстрирует высокую точность определения долей атомов платины в составе наночастиц различных архитектур в агломерате с коэффициентом детерминации R2 более 0.98 и стандартным отклонением ошибок 2.6 %. Апробация обученной модели на экспериментальных данных показала ее высокую обобщающую способность, что свидетельствует о перспективности применения такого подхода к определению эффективности расхода платины при синтезе платиносодержащих катализаторов на основе наночастиц.

Финансирование работы

Работа выполнена при поддержке Российского научного фонда (проект 23-21-00526 https://rscf.ru/project/23-21-00526/) в Южном федеральном университете.

Конфликт интересов

Авторы заявляют, что у них нет конфликта интересов.

Список литературы

[1] K. Kodama, T. Nagai, A. Kuwaki, R. Jinnouchi, Y. Morimoto. Nat. Nanotechnol., **16** (2), 140 (2021). DOI: 10.1038/s41565-020-00824-w

[2] A.A. Alekseenko, A.S. Pavlets, S.V. Belenov, O.I. Safronenko, I.V. Pankov, V.E. Guterman. Appl. Surf. Sci., **595**, 153533 (2022). DOI: 10.1016/j.apsusc.2022.153533

[3] S. Hussain, H. Erikson, N. Kongi, A. Sarapuu, J. Solla-Gullón, G. Maia, A.M. Kannan, N. Alonso-Vante, K. Tammeveski. Int. J. Hydrogen Energy, **45** (56), 31775 (2020). DOI: 10.1016/j.ijhydene.2020.08.215

[4] A. Hrnjic, A.R. Kamšek, A. Pavlišič, M. Šala, M. Bele, L. Moriau, M. Gatalo, F. Ruiz-Zepeda, P. Jovanovič, N. Hodnik. Electrochim. Acta, **388**, 138513 (2021). DOI: 10.1016/j.electacta.2021.138513

[5] S. Zaman, L. Huang, A.I. Douka, H. Yang, B. You, B.Y. Xia. Angew. Chemie Int. Ed., **60** (33), 17832 (2021). DOI: 10.1002/anie.202016977

[6] M. Shao, Q. Chang, J.-P. Dodelet, R. Chenitz. Chem. Rev., **116** (6), 3594 (2016). DOI: 10.1021/acs.chemrev.5b00462

[7] W. Yan, D. Zhang, Q. Zhang, Y. Sun, S. Zhang, F. Du, X. Jin. J. Energy Chem., **64**, 583 (2022). DOI: 10.1016/j.jechem.2021.05.003

[8] M. Heinz, V.V. Srabionyan, L.A. Avakyan, A.L. Bugaev, A.V. Skidanenko, S.Y. Kaptelinin, J. Ihlemann, J. Meinertz, C. Patzig, M. Dubiel, L.A. Bugaev. J. Alloys Compd., **767**, 1253 (2018). DOI: 10.1016/j.jallcom.2018.07.183

[9] S. Belenov, A. Alekseenko, A. Pavlets, A. Nevelskaya, M. Danilenko. Catalysts, **12** (6), 638 (2022). DOI: 10.3390/catal12060638

[10] A.A. Alekseenko, V.E. Guterman, S.V. Belenov, V.S. Menshikov, N.Y. Tabachkova, O.I. Safronenko, E.A. Moguchikh. Int. J. Hydrogen Energy, **43** (7), 3676 (2018). DOI: 10.1016/j.ijhydene.2017.12.143

[11] С.В. Беленов, В.Е. Гутерман, Н.Ю. Табачкова, Е.А. Могучих, А.А. Алексеенко, В.А. Волочаев, Н.М. Новиковский. Электрохимия, **54** (11), 944 (2018). DOI: 10.1134/S042485701813008X [S.V. Belenov, V.E. Guterman, N.Y. Tabachkova, E.A. Moguchikh, A.A. Alekseenko, V.A. Volochaev, N.M. Novikovskiy. Russ. J. Electrochem., **54** (12), 1209 (2018). DOI: 10.1134/S1023193518130062]
[12] K. Boldt, S. Bartlett, N. Kirkwood, B. Johannessen. Nano Lett., **20** (2), 1009 (2020). DOI: 10.1021/acs.nanolett.9b04143
[13] X. Lyu, Y. Jia, X. Mao, D. Li, G. Li, L. Zhuang, X. Wang, D. Yang, Q. Wang, A. Du, X. Yao. Adv. Mater., **32** (32), 2003493 (2020). DOI: 10.1002/adma.202003493
[14] S. Mourdikoudis, R.M. Pallares, N.T.K. Thanh. Nanoscale, **10** (27), 12871 (2018). DOI: 10.1039/C8NR02278J
[15] L.J. Moriau, A. Hrnjić, A. Pavlišiš, A.R. Kamšek, U. Petek, F. Ruiz-Zepeda, M. Šala, L. Pavko, V.S. Šelih, M. Bele, P. Jovanovič, M. Gatalo, N. Hodnik. IScience, **24** (2), 102102 (2021). DOI: 10.1016/j.isci.2021.102102
[16] J. Timoshenko, D. Lu, Y. Lin, A.I. Frenkel. J. Phys. Chem. Lett., **8** (20), 5091 (2017). DOI: 10.1021/acs.jpclett.7b02364
[17] J. Timoshenko, A.I. Frenkel. ACS Catal., **9** (11), 10192 (2019). DOI: 10.1021/acscatal.9b03599
[18] L. Avakyan, D. Tolchina, V. Barkovski, S. Belenov, A. Alekseenko, A. Shaginyan, V. Srabionyan, V. Guterman, L. Bugaev. Comput. Mater. Sci., **208**, 111326 (2022). DOI: 10.1016/j.commatsci.2022.111326
[19] E. Collet, M. Buron, H. Cailleau, M. Lorenc, M. Servol, P. Rabiller, B. Toudic. X-ray diffraction for material science, in: UVX 2008–9e Colloq. Sur Les Sources Cohérentes Incohérentes UV, VUV X Appl. Développements Récents, EDP Sciences, Les Ulis, France, 2009, p. 21. DOI: 10.1051/uvx/2009005
[20] L.A. Bugaev, L.A. Avakyan, V. V. Srabionyan, A.L. Bugaev. Phys. Rev. B, **82** (6), 064204 (2010). DOI: 10.1103/PhysRevB.82.064204
[21] D.C. Koningsberger, B.L. Mojet, G.E. van Dorssen, D.E. Ramaker. Top. Catal., **10** (3/4), 143 (2000). DOI: 10.1023/A:1019105310221
[22] J.A. van Bokhoven, C. Lamberti. X-Ray Absorption and X-Ray Emission Spectroscopy (John Wiley & Sons, Ltd, Chichester, UK, 2016), DOI: 10.1002/9781118844243
[23] R. Wang, H. Wang, F. Luo, S. Liao. Electrochem. Energy Rev., **1** (3), 324 (2018). DOI: 10.1007/s41918-018-0013-0
[24] В.В. Китов. Статистика и экономика, **4**, 22 (2016). DOI: 10.21686/2500-3925-2016-4-22-26 [V.V. Kitov. Stat. Econ., **4**, 22 (2016). DOI: 10.21686/2500-3925-2016-4-22-26]
[25] L. Prokhorenkova, G. Gusev, A. Vorobev, A.V. Dorogush, A. Gulin. (2017). http://arxiv.org/abs/1706.09516
[26] Электронный ресурс. Режим доступа: https://catboost.ai/en/docs/ [Electronic source. Available at: https://catboost.ai/en/docs/]

Приложение 1

Вывод формулы (1)

Уравнение связывает координационные числа (КЧ) металла, получаемые из EXAFS, с координационными числами в материале с учетом того, что атомы металла могут находиться в двух состояниях, принадлежащих двум компонентам: НЧ, далее — NP, и оксиду, далее — Ox.

Координационные числа обозначаются N, принадлежность к фазе — верхний индекс в скобках $N^{(\mathrm{NP})}$, $N^{(\mathrm{Ox})}$. В нижнем индексе информация о паре атомов M–X, центральный атом M — атом окружения X (M = Cu, Pt, X = Cu, Pt, O). Координационное число не является аддитивной величиной по состояниям атомов, но его можно выразить через аддитивные — число связей $n_{\mathrm{M\text{-}X}}$ и число атомов n_{M}, как

$$N_{\mathrm{M}-X} = \frac{n_{\mathrm{M\text{-}X}}}{n_{\mathrm{M}}}.$$

Запишем координационные числа металлов с учетом того, что кислород может быть связан с атомами металла как оксида, так и НЧ:

$$N^{(\mathrm{NP})}_{\mathrm{M\text{-}M}} = \frac{n^{(\mathrm{NP})}_{\mathrm{M\text{-}M}}}{n^{(\mathrm{NP})}_{\mathrm{M}}}, \quad N^{(\mathrm{Ox})}_{\mathrm{M\text{-}O}} = \frac{n^{(\mathrm{Ox})}_{\mathrm{M\text{-}O}}}{n^{(\mathrm{Ox})}_{\mathrm{M}}} \quad \text{и} \quad N^{(\mathrm{NP})}_{\mathrm{M\text{-}O}} = \frac{n^{(\mathrm{NP})}_{\mathrm{M\text{-}O}}}{n^{(\mathrm{NP})}_{\mathrm{M}}}.$$

Число связей с кислородом на поверхности НЧ $n^{(\mathrm{NP})}_{\mathrm{M\text{-}O}}$ зависит от числа атомов металла на поверхности НЧ и ожидается малой величиной для крупных НЧ, с малой долей поверхностных атомов.

Соответствующее координационное число $N^{(\mathrm{NP})}_{\mathrm{M\text{-}O}}$ также должно быть малым.

В EXAFS происходит усреднение по атомам металлов всех состояний, т. е.

$$N^{\mathrm{EXAFS}}_{\mathrm{M\text{-}M}} = \frac{n^{(\mathrm{NP})}_{\mathrm{M\text{-}M}}}{n^{(\mathrm{NP})}_{\mathrm{M}} + n^{(\mathrm{Ox})}_{\mathrm{M}}} \quad \text{и} \quad N^{\mathrm{EXAFS}}_{\mathrm{M\text{-}O}} = \frac{n^{(\mathrm{NP})}_{\mathrm{M\text{-}O}} + n^{(\mathrm{Ox})}_{\mathrm{M\text{-}O}}}{n^{(\mathrm{NP})}_{\mathrm{M}} + n^{(\mathrm{Ox})}_{\mathrm{M}}}.$$

Перейдем от числа связей к координационному числу:

$$N^{\mathrm{EXAFS}}_{\mathrm{M\text{-}M}} = \frac{n^{(\mathrm{NP})}_{\mathrm{M}} \cdot N^{(\mathrm{NP})}_{\mathrm{M\text{-}M}}}{n^{(\mathrm{NP})}_{\mathrm{M}} + n^{(\mathrm{Ox})}_{\mathrm{M}}}$$

$$\text{и} \quad N^{\mathrm{EXAFS}}_{\mathrm{M\text{-}O}} = \frac{n^{(\mathrm{NP})}_{\mathrm{M}} \cdot N^{(\mathrm{NP})}_{\mathrm{M\text{-}O}} + n^{(\mathrm{Ox})}_{\mathrm{M}} \cdot N^{(\mathrm{Ox})}_{\mathrm{M\text{-}O}}}{n^{(\mathrm{NP})}_{\mathrm{M}} + n^{(\mathrm{Ox})}_{\mathrm{M}}}$$

и обозначим соотношение количеств атомов (концентраций): $\xi = \frac{n^{(\mathrm{NP})}_{\mathrm{M}}}{n^{(\mathrm{NP})}_{\mathrm{M}} + n^{(\mathrm{Ox})}_{\mathrm{M}}}$. Величина ξ имеет смысл доли атомов платины в составе НЧ. При этом $1 - \xi = \frac{n^{(\mathrm{Ox})}_{\mathrm{M}}}{n^{(\mathrm{NP})}_{\mathrm{M}} + n^{(\mathrm{Ox})}_{\mathrm{M}}}$ — доля атомов платины в оксиде. Тогда координационные числа, получаемые из EXAFS:

$$N^{\mathrm{EXAFS}}_{\mathrm{M\text{-}M}} = \xi \cdot N^{(\mathrm{NP})}_{\mathrm{M\text{-}M}} \tag{П1}$$

$$N^{\mathrm{EXAFS}}_{\mathrm{M\text{-}O}} = \xi \cdot N^{(\mathrm{NP})}_{\mathrm{M\text{-}O}} + (1 - \xi) \cdot N^{(\mathrm{Ox})}_{\mathrm{M\text{-}O}}. \tag{П2}$$

Из равенства (П2) находим ξ:

$$\xi = \frac{N^{(\mathrm{Ox})}_{\mathrm{M\text{-}O}} - N^{\mathrm{EXAFS}}_{\mathrm{M\text{-}O}}}{N^{(\mathrm{Ox})}_{\mathrm{M\text{-}O}} - N^{(\mathrm{NP})}_{\mathrm{M\text{-}O}}},$$

тогда из (П1) получаем связь между координационными числами атомов металла в НЧ и получаемыми из EXAFS:

$$N_{\text{M-M}}^{\text{EXAFS}} = \frac{N_{\text{M-O}}^{(\text{Ox})} - N_{\text{M-O}}^{\text{EXAFS}}}{N_{\text{M-O}}^{(\text{Ox})} - N_{\text{M-O}}^{(\text{NP})}} N_{\text{M-M}}^{(\text{NP})}.$$

И, наконец, пренебрегая окислением атомов на поверхности НЧ $N_{\text{M-O}}^{(\text{Ox})} \gg N_{\text{M-O}}^{(\text{NP})}$, получим:

$$N_{\text{M-M}}^{(\text{NP})} = \frac{N_{\text{M-O}}^{(\text{Ox})}}{N_{\text{M-O}}^{(\text{Ox})} - N_{\text{M-O}}^{\text{EXAFS}}} N_{\text{M-M}}^{\text{EXAFS}}.$$

Приложение 2

Подбор гиперпараметров используемых моделей

Для модели CatBoostRegressor был выполнен подбор следующих гиперпараметров в пределах:

- iterations: [50, 100, 500, 1000, 1500],
- depth: [5, 6, 7, 8, 10, 12],
- learning_rate: [0.001, 0.01, 0.05, 0.1, 0.5],
- l2_leaf_reg: [1, 3, 5, 10].

С помощью метода поиска по сетке были найдены следующие значения гиперпараметров: iterations = 100, depth = 7, learning_rate = 0.5, l2_leaf_reg = 5.

Для настройки гиперпараметров нейронной сети был применен комбинированный подход, включающий грубый подбор и поэтапное усложнение архитектуры. Мы начинали с простых конфигураций и постепенно увеличивали сложность модели, добавляя слой за слоем и корректируя другие параметры, пока не наблюдали признаки переобучения на небольшой подвыборке тренировочных данных (20 %). Это позволило определить оптимальный баланс между сложностью модели и ее обобщающей способностью.

Кроме того, были протестированы различные типы архитектур, включая многослойный персептрон (MLP) и сверточную нейронную сеть (CNN). Сравнительный анализ показал, что многослойный персептрон демонстрирует более высокую эффективность для решения поставленной задачи, что подтверждалось метриками качества на валидационных и тестовых данных.