

Создание нейронной сети для анализа локальной атомной структуры

Э. Ф. Хаметова^{1,2*} О. Р. Бакиева^{2†}

*Удмуртский государственный университет, Институт математики
информационных технологий и физики, кафедра общей физики*

Россия, 426034, Удмуртская республика, Ижевск, ул. Университетская, 1

*Удмуртский федеральный исследовательский центр Уральского отделения Российской Академии Наук,
Физико-технический институт*

Россия, 426000, Ижевск, ул. Кирова, 132

(Поступила в редакцию 17.05.2022; подписана в печать 10.06.2022)

В данной работе рассматривается возможность исследования локальной атомной структуры вещества методом анализа протяженных тонких структур спектров энергетических потерь электронов (EXELFS — Extended Electron Energy Loss Fine Structure) с применением технологий нейронных сетей для определения параметров структуры: парциальных координационных чисел, длин химической связи и параметров тепловой дисперсии. В качестве модельных объектов были выбраны экспериментальные данные EXELFS спектров однокомпонентных систем, полученные в режиме на просвет. На первом этапе работы в качестве определяемого параметра было выбрано межатомное расстояние, поскольку для получения данной характеристики исходным экспериментальным данным требуется минимальная обработка. Рассмотрены различные методы обучения и виды нейронных сетей: обучение с учителем, обучение без учителя и обучение с подкреплением. Проведен анализ работ исследователей в области спектроскопии, решающих различные задачи с применением методов машинного обучения. Показаны результаты работы дополнительного модуля, способного анализировать спектры однокомпонентных систем.

PACS: 07.05.Mh.

УДК: 004.942.

Ключевые слова: EXELFS-спектроскопия, локальная атомная структура, машинное обучение, модуль, однокомпонентные системы.

ВВЕДЕНИЕ

Нейронные сети — одно из перспективных направлений искусственного интеллекта, которое на сегодняшний день активно используется во многих сферах деятельности человека [1, 2]. Они применяются для решения сложных задач (например, классификация, прогнозирование, распознавание), которые требуют аналитических вычислений подобных тем, которые делает человеческий мозг. Популярность нейронных сетей обусловлена рядом особенностей. К их числу относится параллельная обработка информации всеми звеньями сети, что значительно ускоряет процесс анализа данных. Важной особенностью является способность к обобщению полученной информации и дальнейшее применение накопленных «знаний» при обработке новых данных, иными словами — обучение. Последнее является главным отличием нейронных сетей от обычных алгоритмов. Для работы нейронной сети необходимо грамотно выбрать ее архитектуру и передать ей правильную выборку, по которой она сможет обучиться.

В настоящее время самые актуальные разработки и инновации сконцентрированы вокруг материалов с низкой размерностью — это и сложная наноструктурированная керамика, и нанокомпозитные материалы для фотоники, и топологические сверхпроводни-

ки для нового поколения квантовых компьютеров [3]. Естественным образом вместе с появлением новых материалов встает вопрос об их корректной диагностике, поскольку часто высокие функциональные свойства достигаются за счет строго определённой локальной координации атомов.

Классическим методом анализа локальной атомной структуры является EXAFS (Extended X-ray Absorption Fine Structure) спектроскопия. Структура EXAFS спектра формируется в результате дипольного процесса когерентного рассеяния фотоэлектрона на локальном окружении [4]. В случае исследования локального окружения легких элементов удобно использовать для возбуждения внутреннего уровня атома поток электронов, т.е. метод EXELFS (Extended Electron Energy Loss Fine Structure) спектроскопии [5–7]. В данном методе используется поток моноэнергетичных электронов и происходит регистрация потерь энергии электронов на локальном окружении возбуждаемого атома. Таким образом, протяженная тонкая осциллирующая структура формируется в результате когерентного рассеяния вторичного электрона на локальном окружении возбуждаемого атома и несет в себе информацию о параметрах локального атомного окружения. Период осцилляций тонкой структуры 15–20 эВ, а их протяженность от 50 эВ до нескольких сотен эВ, в зависимости от сорта возбуждаемого атома. Данные спектры относятся к EXAFS-подобным и могут быть получены на просвечивающем микроскопе в лабораторных условиях.

* elinaphanilevna851@gmail.com

† ftiran@mail.ru

На рис. 1 представлен экспериментальный EXELFS спектр в шкале энергетических потерь электронов, где можно наблюдать край возбуждения атома Ti, за которым следует тонкая осциллирующая структура. Данный спектр был взят из открытой базы данных EELS DB [8].

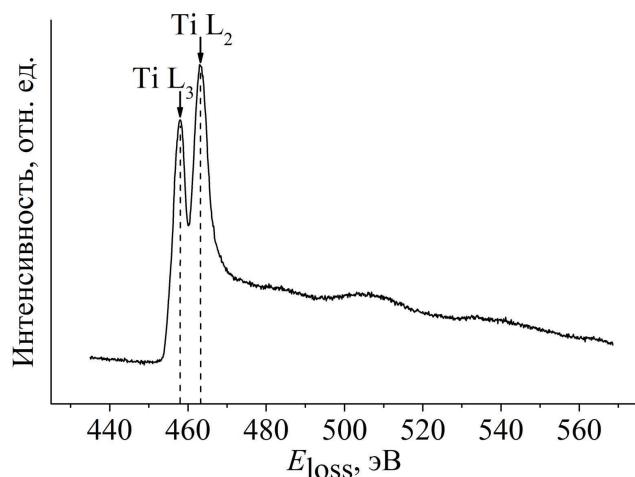


Рис. 1. Экспериментальный спектр энергетических потерь электронов Ti [8]: L3 — внутренний уровень титана с энергией связи 455 эВ, L2 — внутренний уровень атома титана с энергией связи 461 эВ

Анализ данных спектров дает информацию о том, на каком расстоянии расположены атомы, какого сорта и в каком количестве. Обычно обработка экспериментальных данных состоит из нескольких этапов. Первый этап — вычитание фона. Вычитаемый фон содержит процессы, не относящиеся к когерентному рассеянию. Второй этап — нормировка. Возможно, проводить анализ и ненормированных спектров, но в таком случае извлекаемая информация будет полуколичественной — т.е. возможно определить межатомные расстояния, но не координатные числа. Поэтому для корректной обработки экспериментальных данных необходимо нормировать экспериментальный спектр на функцию, не зависящую от структуры исследуемого вещества. Третий этап — Фурье-преобразование, в результате которого получают набор Фурье-пику. Положение пика говорит о межатомном расстоянии, а взвешенная площадь — о координационном числе.

Обработка EXELFS спектров является сложной задачей, поскольку для проведения структурных исследований экспериментальных данных необходимо преодолеть ряд проблем. Во-первых, нужно уметь обрабатывать данные эксперимента, не внося дополнительные погрешности и шум, то есть вычитаемый фон должен описываться гладкой неосциллирующей функцией, чтобы не вносить какие-либо искажения в экспериментальные данные. Во-вторых, возникает проблема E_0 . Волновой вектор электрона обладает нелинейной зависимостью от выбора потенциала ионизации E_0 .

И до сих пор нет однозначного способа его определения. Вследствие чего возникает дополнительная ошибка при определении межатомных расстояний. В-третьих, выделяемого диапазона может быть недостаточно для корректной обработки методом Фурье-преобразования, что приводит к сильному уширению Фурье-пику и искажению результатов. В-четвертых, необходимо получить априорную информацию об исследуемом объекте. Она необходима для построения модели корректно описывающей исследуемую структуру реального объекта [4]. Комплекс этих проблем можно избежать при использовании нейронной сети, способной обрабатывать экспериментальные спектры энергетических потерь электронов.

Данная работа является одним из этапов построения нейронной сети для количественного анализа EXELFS спектров. Цель работы — построение модуля для определения химического элемента, к которому принадлежит экспериментальный спектр, и соответствующего края возбуждения. Поэтому создание нейронной сети, способной проводить анализ EXELFS спектров значительно облегчит процесс обработки экспериментальных данных и получения информации о локальной атомной структуре.

1. ПРИМЕНЕНИЕ НЕЙРОННЫХ СЕТЕЙ В ЭЛЕКТРОННОЙ СПЕКТРОСКОПИИ

На данный момент существует несколько направлений в спектроскопии, в которых используются нейронные сети. В связи с этим был проведен анализ литературы, и ниже приведены примеры работ, наиболее близко относящиеся к рассматриваемой проблематике.

Авторы работы [9] представили новую независимую от модели стратегию параметризации и вычитания пика упруго-отраженных электронов, который доминирует в области с низкими потерями. Используемая стратегия основана на методах машинного обучения и обеспечивает точную оценку неопределенностей, связанных как с входными данными, так и с самой процедурой анализа, которая затем может быть распространена на физические прогнозы без каких-либо приближений. Для оценки использовался метод реплик Монте-Карло, а оптимизация работы нейронной сети осуществлялась с помощью стохастического градиентного спуска в сочетании с обратным распространением ошибки. Благодаря этому можно в полной мере использовать ценную физическую информацию, содержащуюся в области спектров с низкими потерями. Таким образом, решается важная задача вычитания пика упруго-отраженных электронов, однако работы ведутся в области начала координат в шкале потерь энергии электронов и данное решение не дает напрямую информации о локальной атомной структуре.

Авторы работы [10] создали нейронную сеть для непосредственного прогнозирования функции радиального распределения из ближней краевой области спек-

тра рентгеновского поглощения (XANES), что позволило преодолеть ряд трудностей, таких как: слабые сигналы, необходимость широкого диапазона энергий и подгонка параметров в эталонных образцах. В качестве набора данных в работе использовались модельные спектры, которые не содержат случайных шумов или ошибок. Для оптимизации параметров обучения использовалось обратное распространение ошибок по схеме Адама, минимизируя среднеквадратичные ошибки. С помощью созданной нейронной сети успешно были предсказаны функции радиального распределения оксидов кремния из данных XANES. Таким образом, авторы предлагают метод получения информации о локальной атомной структуре основанный на обучении на модельных спектрах, которые могут существенно отличаться от экспериментальных наличием шумов и нелинейного фона.

Авторы работы [11] использовали технологии машинного обучения для анализа геометрии молекул. Представленная нейронная сеть способна на основе спектра XANES прогнозировать структурные параметры наноструктуры. В ходе работы были исследованы два подхода к применению машинного обучения: прямой (модель обучается прогнозировать структурные параметры непосредственно по спектру XANES) и обратный (модель используется для аппроксимации спектра как функции структурных параметров, после чего некоторым алгоритмом подбираются параметры, соответствующие спектру, наиболее близкому к экспериментальному). Были определены параметры кристаллической структуры CPO-27-Ni после адсорбции газа. Полученные результаты хорошо согласуются с экспериментальными значениями, полученными дифракцией рентгеновских лучей методом Ритвельда. Таким образом, создана нейронная сеть, способная прогнозировать параметры кристаллической структуры по рентгеновским спектрам области XANES. В электронной спектроскопии данный метод можно также применить для обработки сигнала в околокраевой области, однако в данной работе мы рассматриваем диапазон экспериментальных данных до 500 эВ после края энергетических потерь электрона.

Из рассмотренных выше работ можно выделить следующее. В последние несколько лет в области спектроскопии активно начинают применяться нейронные сети. Однако в работах [9–11] рассматривается только около краевая область низких потерь и выход на информацию о локальной атомной структуре проходит через промежуточный этап. В настоящей работе исследуется область за краем характеристических потерь энергии электронов и предлагается с помощью нейронной сети получать параметры локальной атомной структуры без промежуточного этапа, то есть непосредственно из диапазона, содержащего эти данные.

2. МЕТОДЫ ОБУЧЕНИЯ НЕЙРОННЫХ СЕТЕЙ

Для того чтобы приступить к написанию нейронной сети, были проанализированы все методы обучения и модели нейронных сетей. Этот этап необходим для выбора наиболее подходящего метода обучения и архитектуры нейронной сети. На сегодняшний день существует три типа методов обучения искусственных нейронных сетей: с учителем, без учителя и с подкреплением [12, 13].

Обучение с учителем. Процесс обучения с учителем предполагает, что для каждого из множества входных векторов существует целевой вектор, определяющий желаемый выход. Иначе говоря, для каждого прецедента задается пара — «исходные данные — правильное решение», то есть данные являются размеченными. Такая пара называется обучающей выборкой. В данном случае алгоритм обучения основывается на минимизации ошибки сети и корректировки весовых коэффициентов. При этом важно учитывать, что размер обучающей выборки должен быть достаточно большим. В зависимости от поставленной задачи и её сложности обучающая выборка может содержать от 500 объектов. Согласно [13] для приемлемого качества работы нейронной сети необходимо наличие около 5000 помеченных примеров. Обучение с учителем применяют, например, для классификации различных объектов (цифры, цветы, животные), распознавания звуков, прогнозирования временных рядов.

Обучение без учителя. Процесс обучения без учителя не требует наличия целевого вектора выходов, т.е. нет необходимости в сравнении получаемых результатов с некими эталонными ответами. Иными словами для каждого прецедента задаются только «исходные данные». Такая обучающая выборка называется неразмеченной. Алгоритм обучения в данном случае корректирует весовые коэффициенты таким образом, чтобы сеть могла из схожих по некоторому принципу входных данных выдать результат, обнаруживающий другие взаимосвязи и закономерности между этими данными. Для такого алгоритма не нужно обирать большую размеченную базу данных. Обучение без учителя применяют для кластеризации объектов, языковых моделей, статистических моделей.

С подкреплением. Обучение с подкреплением — это комбинация описанных выше подходов к обучению. В данном случае сеть обучается на неразмеченной обучающей выборке. Однако есть необходимость в обратном отклике на действия нейронной сети, то есть показать приводят ли действия к правильному ответу или нет.

Анализ, проведенный по видам обучения показал, что для разработки нейронной сети, способной обрабатывать электронные спектры и определять параметры локальной атомной структуры, лучше всего подходит метод обучения с подкреплением. Такой выбор объясняется отсутствием необходимости в размеченной обучающей выборке, одновременно с возможностью контроля и оценки работы нейронной сети.

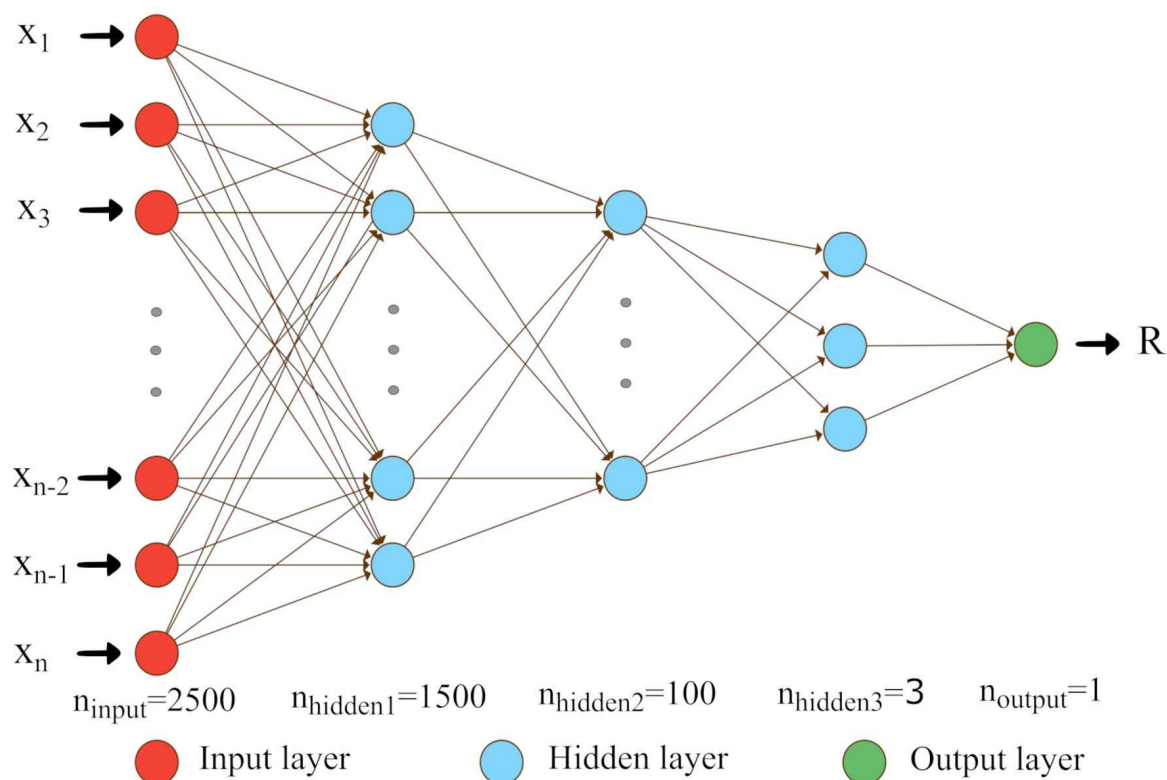


Рис. 2. Схема архитектуры нейронной сети: n — количество точек в спектре (входные данные), R — межатомное расстояние (выходной параметр), n_{input} — количество узлов во входном слое, $n_{hidden1}$ — количество узлов в первом скрытом слое, $n_{hidden2}$ — количество узлов во втором скрытом слое, $n_{hidden3}$ — количество узлов в третьем скрытом слое, n_{output} — количество узлов в выходном слое

Таким образом, на первом этапе планируется обучить нейронную сеть обрабатывать экспериментальные данные, полученные в режиме на просвет. При обучении определяемым параметром является межатомное расстояние, поскольку для определения этого параметра экспериментальным данным не требуется нормировка.

3. СОСТОЯНИЕ РАБОТЫ НА СЕГОДНЯШНИЙ ДЕНЬ

В данной работе нейронная сеть создавалась без использования дополнительных фреймворков специально для нейронных сетей методом обучения с учителем и имеет следующую архитектуру, представленную на рис. 2: входной слой (2500 узлов), затем 3 скрытых (1500, 100 и 3 узла соответственно) и выходной слой (1 узел).

Коэффициент обучения равнялся 0.1, а в качестве функции активации использовалась сигмоидальная функция, везде кроме последнего слоя. Функция активации определяет выходное значение нейрона в зависимости от результата взвешенной суммы входов и порогового значения. Сигмоида — гладкая нелиней-

ная функция, подобная ступенчатой функции (рис. 3).

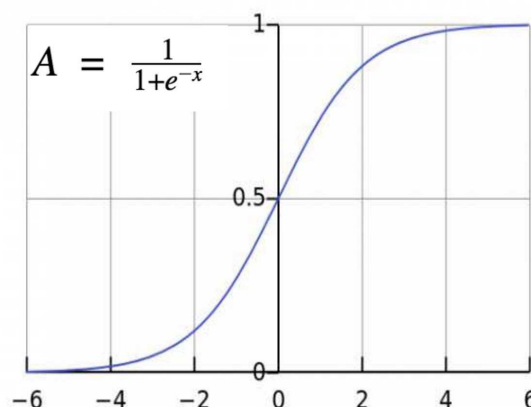


Рис. 3. Аналитическое и графическое представление сигмоиды

В последнем слое не использовалась функция активации, что позволит получить ожидаемый результат из всего множества чисел (от $-\infty$ до $+\infty$), а не из некоторого промежутка (от -1 до 1 или от 0 до 1).

Результат работы данной сети показал, что имеющейся обучающей выборки недостаточно для достижения поставленной цели. Варьирование скрытых слоев и количества узлов не решило проблемы с обучением сети.

Данные результаты привели к следующим выводам:

- либо необходимо увеличить базу размеченных данных для обучения сети;
- либо, возможно, необходимо поменять стратегию обучения сети.

Было решено перейти к методу обучения с подкреплением, а также последовать примеру авторов работы [11] и использовать для создания сети библиотеку TensorFlow с фреймворком Keras. В статье [11] использовалась последовательная архитектура сети со слоем свертки вначале, затем с тремя полносвязанными слоями с функциями активации ReLU. ReLU — нелинейная функция активации, возвращающая значение x , если x положительно, и 0 в противном случае (рис. 4).

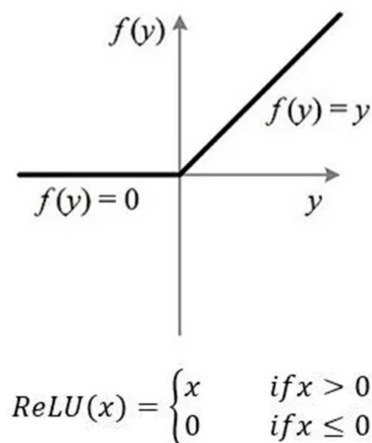


Рис. 4. Аналитическое и графическое представление функции активации ReLU

Контроль качества аппроксимации функции и оценка погрешности предсказываемых значений проводилась с помощью кросс-валидации. Кросс-валидация (или перекрестная проверка) — метод оценки аналитической модели и ее поведения на независимых данных. В ходе оценки модели имеющиеся данные разбиваются на k частей. После чего обучение модели производится на $k - 1$ частях данных, а оставшаяся часть используется в качестве тестовой. Данная процедура повторяется k раз. Таким образом, каждая из k частей данных используется для тестирования. В результате этого оценка эффективности выбранной модели получается с наиболее равномерным использованием имеющихся данных. Этот метод повышает надежность модели, удерживая данные вне процесса обучения. Кроме повышения производительности на многих не учитываемых наблюдениях, в средах с ограниченными данными он может быть эффективным инструментом для

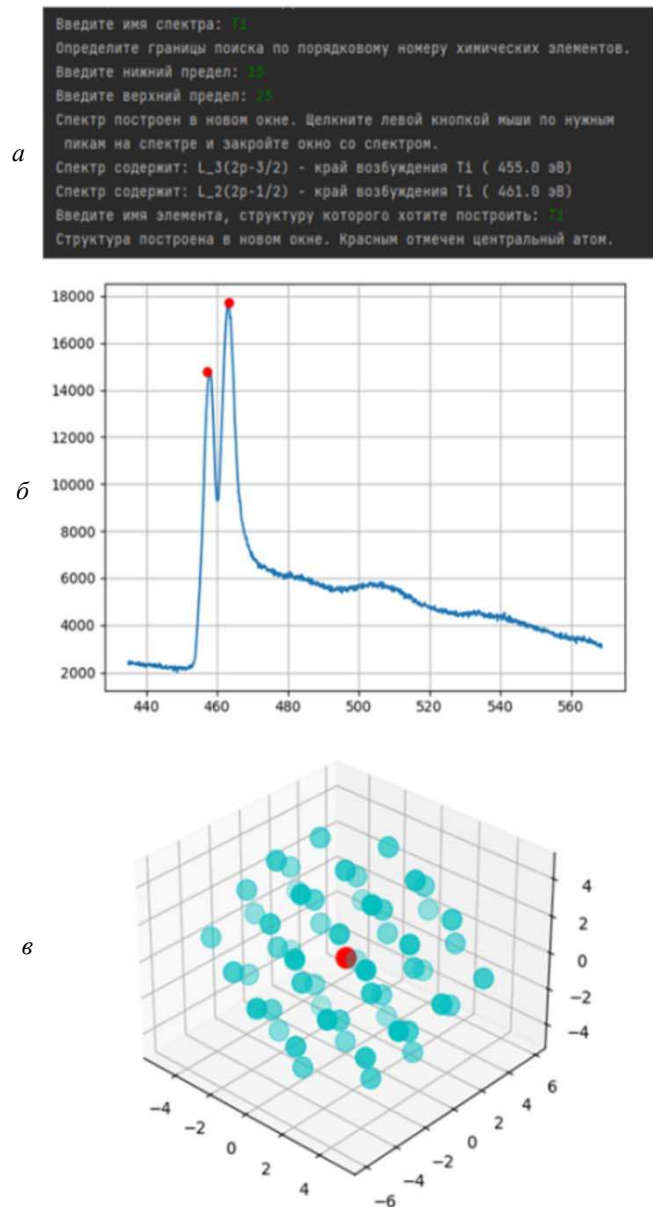


Рис. 5. Работа модуля на примере экспериментальных данных титана: а — окно ввода входных параметров (зеленый шрифт) и вывод результата работы программы (определение краев возбуждения); б — пример ввода положений краёв L_2 и L_3 титана (пример выделения пиков, интересующих оператора (в нашем случае получаем L_2 и L_3 края характеристических потерь титана)); в — предлагаемая программой структура (красным отмечен центральный атом)

обучения моделей с меньшим набором данных, что актуально для рассматриваемой задачи.

Для оптимизации весов использовался итерационный метод SGD (Stochastic Gradient Descent — стохастический градиентный спуск) в сочетании с обратным распространением ошибки, минимизирующим среднеквадратичную ошибку путем корректировки весов сети. Метод обратного распространения ошибки

используется для эффективного обучения сети с помощью правила дифференцирования сложной функции. То есть после каждого прохода по нейронной сети выполняется проход в обратную сторону и регулируются параметры модели (веса).

В рамках поставленной цели разработан и полностью готов к использованию модуль для нейронной сети, в котором заложена необходимая для дальнейшей обработки информация: сведения об энергии связи [14] и структурные данные, взятые из базы данных модуля ATOMS [15]. Модуль на входе принимает экспериментальный спектр энергетических потерь электронов однокомпонентной системы. Далее оператору (с помощью внешних устройств ввода — клавиатура/мышь) необходимо указать в экспериментальных данных положение края характеристических потерь (рис. 5, б), обычно это максимально интенсивный пик в спектре. Программный модуль по указанному положению пика определяет, к какому химическому элементу принадлежит спектр и соответствующий край возбуждения, а также строит предполагаемую атомную структуру исследуемого вещества (рис. 5, в). Результат работы программы представлен на рис. 5.

ЗАКЛЮЧЕНИЕ

На данном этапе проведен анализ имеющихся работ по созданию и использованию нейронных сетей с целью получения информации из электронных спектров. Проведенный литературный обзор показал, что к настоящему моменту не разработана нейронная сеть для

получения информации о параметрах локальной атомной структуры по экспериментальным данным протяженной тонкой структуры спектров энергетических потерь электронов. Таким образом, решаемая в данной работе задача является актуальной. Из рассмотренных в данной работе принципов построения нейронных сетей была выбрана наиболее подходящая архитектура сети: полносвязанные слои с сигмоидальной функцией активации во всех слоях, кроме выходного слоя, где активационная функция — ReLU. Было решено при построении сети использовать библиотеку TensorFlow с фреймворком Keras.

На данном этапе работы создан модуль, анализирующий спектры однокомпонентных систем, и позволяющий определять по положению пика в шкале потерь энергии, к какому химическому элементу принадлежит спектр и соответствующий край возбуждения, а также предлагается атомная структура исследуемого вещества, используя данные о структуре из модуля ATOMS. Следующим этапом планируется расширить диапазон экспериментальных данных для обучения нейронной сети до многокомпонентных систем и расширить функциональные свойства программного модуля до получения функции радиального распределения атомов.

Работа выполнена в рамках государственного задания Министерства науки и высшего образования РФ №121030100002-0. Исследования выполняли с использованием оборудования ЦКП «Центр физических и физико-химических методов анализа, исследования свойств и характеристик поверхности, наноструктур, материалов и изделий» УдмФИЦ УрО РАН.

-
- [1] Абрикосов И. А., Сон Э.Е., Мухамедов Б.О., Хван А.В. // Теплофизика высоких температур. 2020. **58**. С. 915.
 - [2] Schmeide K. et al. // Science of the Total Environment. 2021. **770**.
 - [3] Yangjin Dong et al. // International Journal of Biological Macromolecules. 2021. **190**. P. 693.
 - [4] Кочубей Д.И., Бабанов Ю.А., Замараев К.И. и др. Рентгеноспектральный метод изучения структуры аморфных тел: EXAFS-спектроскопия. Наука, 1988.
 - [5] Crescenzi M. De. // Crit. Rev. Sol St. Mater. Sci. 1989. **15**.
 - [6] Ruts Yu.V., Guy D.E., Surnin D.V., Grebennikov V.I. Secondary Electron Fine Structure — A Method of Local Atomic Structure Characterization. In: H.S. Nalwa (Ed.) Handbook of Surfaces and Interfaces of Materials. Academic Press. 2001. **2**.
 - [7] Guy D., Bakieva O., Grebennikov V., Surnin D., Ruts Y. // J. El. Spectr. Rel. Phen., 2010. **182**.
 - [8] Philip Ewels, Thierry Sikora, Virginie Serin, Chris P. Ewels and Luc Lajaunie. [A Complete Overhaul of the Electron Energy-Loss Spectroscopy and X-Ray Absorption Spectroscopy Database: eelsdb.eu.](#) // [Microscopy and Microanalysis, available on CJO2016.](#)
 - [9] Roest L., van Heijst S.E. et al. // Ultramic. 2021.
 - [10] Shin Kiyohara, Teruyasu Mizoguchi // J. Phys. Soc. Jpn. 2020. **89**.
 - [11] Гуда С.А., Алгасов А.С. // Вестник РГУПС. 2019. № 2. С. 84 .
 - [12] Гафаров Ф.М., Галимянов А.Ф. Искусственные нейронные сети и их приложения: учеб. пособие. Изд-во Казан. ун-та, 2018.
 - [13] Гудфеллоу Я., Бенжио И., Курвилль А. Глубокое обучение / пер. с англ. А.А. Сликина. М.: ДМК Пресс, 2018.
 - [14] Зигбан К., Нордлинг К., Фальман А. и др. Электронная спектроскопия. Перевод с англ. под редакцией д-ра физ.-матем. наук, проф. И. Б. Боровского. М. «Мир», 1971.
 - [15] Ankudinov A.L., Rehr J.J. Relativistic Spin-dependent X-ray Absorption Theory // UW preprint. 1996.

Neural network for local atomic structure analysis

E.F. Khametova^{1,2,a}, O.R. Bakieva^{2,b}

¹*Udmurt State University, Institute of Mathematics, Information Technology and Physics, Department of General Physics
Udmurt Republic, Izhevsk, 426034, Russia*

²*Udmurt Federal Research Center of the Ural Branch of the Russian Academy of Sciences, Institute of Physics and Technology,
Izhevsk, 426000, Russia*

E-mail: ^aelinaphanilevna851@gmail.com, ^bftiran@mail

In this paper, we consider the possibility of studying the local atomic structure of a substance by analyzing extended fine structures of electron energy loss spectra (EXELFS — Extended Electron Energy Loss Fine Structure) using neural network technologies to determine the structure parameters: partial coordination numbers, chemical bond lengths and thermal dispersion parameters. The experimental data of EXELFS spectra of one-component systems obtained in the transmission mode were chosen as model objects. At the first stage of the work, the interatomic distance was chosen as the parameter to be determined, since minimal processing is required to obtain this characteristic from the initial experimental data. Various learning methods and types of neural networks are considered: supervised learning, unsupervised learning and reinforcement learning. The analysis of the work of researchers in the field of spectroscopy, solving various problems using machine learning methods, is carried out. The results of the operation of an additional module capable of analyzing the spectra of one-component systems are shown.

PACS: 07.05.Mh.

Keywords: EXELFS spectroscopy, local atomic structure, machine learning, module, one-component systems.

Received 17 May 2022.

Сведения об авторах

1. Хаметова Элина Фанилевна — студент; elinaphanilevna851@gmail.com.
2. Бакиева Ольга Ринатовна — канд. физ.-мат. наук, ст. науч. сотрудник; e-mail: ftiran@mail.ru.