

ОБОБЩЕННАЯ МОДЕЛЬ ПРОГНОЗИРОВАНИЯ ГАЗОПРОНИЦАЕМОСТИ СТЕКЛООБРАЗНЫХ ПОЛИМЕРОВ И ОСТАТОЧНЫЕ НЕЙРОННЫЕ СЕТИ КАК ИНСТРУМЕНТ ЕЕ УЛУЧШЕНИЯ

©2024 г. Д. А. Царев^а, В. Е Рыжих^а, Н. А. Белов^а, А. Ю Алентьев^{а, *}

^а Институт нефтехимического синтеза им. А.В. Топчиева Российской академии наук

119991 Москва, Ленинский пр., 29

*e-mail: alentiev@ips.ac.ru

Поступила в редакцию 27.11.2024 г.

После доработки 16.12.2024 г.

Принята к публикации 27.12.2024 г.

В работе продемонстрированы новые возможности для улучшения прогнозирования газотранспортных характеристик стеклообразных полимеров по их химической структуре на основе Базы данных ИНХС РАН. Разработана обобщенная линейная модель для прогнозирования коэффициентов проницаемости любой системы газ–полимер на основе структурных дескрипторов полимера и свойств газа, например, табулированных эффективного кинетического диаметра для молекул газа и эффективного параметра потенциала Леннард-Джонса. Такая модель позволяет существенно расширить массив экспериментальных данных для прогнозирования и применения современных методов машинного обучения. Показана возможность использования малых остаточных нейронных сетей для улучшения качества прогнозов линейных моделей, причем обучение таких нейронных сетей не требует использования серьезных вычислительных мощностей.

DOI: 10.31857/S2308112024040079, EDN: MBVDBC

ВВЕДЕНИЕ

Для прогнозирования физико-химических свойств различных веществ по их химической структуре традиционно используют аддитивные методы [1–6]. В рамках аддитивных методов (методов групповых вкладов) выдвигается предположение о том, что каждый фрагмент химической структуры вещества вносит свой вклад в значение того или иного физико-химического свойства [2, 3]. Первые попытки применения аддитивных методов к полимерным объектам были сделаны в работах [7–10], а значительный прогресс был достигнут в работе Ван-Кревелена [11]: в ней рассматривались возможности расчета и прогнозирования таких свойств полимеров, как плотность, коэффициент термического расширения, теплоемкость, механические свойства и другие. Данное направление впоследствии было расширено в работах J. Vicerano [12], А.А. Аскадского [13, 14], М. Salame [15–17]. Еще в начале 1990-х годов точность прогнозирования транспортных свойств полимеров аддитивными методами была неудовлетворительной [18], однако в конце того же десятилетия появились методы групповых вкладов с достаточно высокой

прогностической способностью [19–21]. Позднее вышел цикл работ, посвященных различным аддитивным подходам к прогнозированию параметров транспорта газов в полимерах [22, 23]. Точность прогнозирования газотранспортных свойств методами, представленными в публикациях [22, 23], развивали в дальнейшем как в работах отечественной группы Ю.П. Ямпольского [24–27], так и в работах зарубежной группы под руководством S.B. Tantekin-Ersolmaz [28].

Тем не менее дальнейшее развитие прогнозирования свойств полимеров аддитивными методами представляется маловероятным, поскольку в настоящее время большинство способов разбиения химической структуры уже было исследовано, а вариабельность состояний стеклообразных полимеров, как и полимеров в целом, приводит к тому, что ошибка прогнозирования таких свойств, как газотранспортные, в целом слабо изменяется.

В настоящее время все большую популярность обретают нейросетевые модели прогнозирования [29–33]. Они используются для решения как научных, так и прикладных задач. Ранее для прогнозирования ряда физико-химических свойств

полимеров использовали методы машинного обучения, а именно, метод, использующий на входе *уникальный набор структурных дескрипторов*¹ для описания каждого полимера (вместо структурной формулы) с обучением посредством методов гребневой регрессии (KRR²) или гауссовой регрессии (GPR³) на основе радиальных базисных функций (RBF⁴) и результаты работы были реализованы в формате онлайн-платформы прогнозирования Polymer Genome [30, 31]. J.W. Barnett и соавт. [32] схожим образом решали задачу поиска структур полимеров, имеющих заданные характеристики газопроницаемости, а в работе [34] метод, использующий уникальный набор дескрипторов, был объединен с формированием синтетической базы полимерных структур (для увеличения количества данных) методом молекулярно-динамического моделирования для прогнозирования доли свободного объема в полимерах. В работе группы С.В. Люлина [29] применяли сверточную нейронную сеть для прогнозирования температуры стеклования полимеров на основе сгенерированной базы данных для полиимидов, сформированной для получения массива данных, достаточного для проведения машинного обучения. Детальный обзор современных нейросетевых архитектур, связанных с химическими приложениями, приведен в работе P. Reiser и соавт. [35].

Лимитирующим фактором для применения нейронных сетей для прогнозирования газотранспортных свойств полимеров является ограниченность массива доступных экспериментальных данных. Как было сказано выше, в ряде работ [29, 34] используют созданные тем или иным расчетным методом массивы данных, чтобы проведение машинного обучения с применением современных нейронных моделей было реализуемо и обосновано. Поскольку в линейных инкрементальных моделях связи структура—свойство обычно рассматривают систему один газ—различные полимеры, количество экспериментальных данных (уравнений) для разных газов может отличаться весьма существенно и, как правило, не превышает 1500, что крайне мало для построения любых нейросетевых моделей. Так как для системы один полимер—разные газы известны хорошо работающие линейные зависимости логарифма коэффициента диффузии от квадрата кинетического диаметра газа d_{eff}^2 и логарифма коэффициента растворимости от параметра потенциала Леннарда-Джонса $(\epsilon/k)_{eff}$ [36–38], существует принципиальная возможность построения обобщенной модели для всех полимеров и газов, т.е. объединения инкрементов, получаемых аддитивными

методами, со свойствами газов. Такая обобщенная модель позволит существенно расширить массив экспериментальных данных для прогнозирования.

Однако этого все еще недостаточно для реализации машинного обучения на современных нейросетевых моделях. Тем не менее классические нейросетевые модели на таких массивах данных уже работоспособны. В случае, если на выходе нейронной сети присутствует только одно числовое значение, классические нейросетевые модели по своей сути ничем не отличаются от обычных нелинейных функций многих переменных, со всеми присущими им достоинствами (гибкость и возможность описать почти любые непротиворечивые данные) и недостатками (такими как сложность подбора параметров модели, а при отсутствии теоретического базиса, как правило, такие модели имеют очень низкую предсказательную способность или — в терминах нейронных сетей — являются переобученными). Хорошая предсказательная способность может быть достигнута только на огромных массивах экспериментальных данных (50000 экспериментальных точек и более [39]), что существенно ограничивает применимость нейронных сетей в современных научных изысканиях. В связи с этим в настоящей работе применяются классические нейронные сети как вспомогательный инструмент для уменьшения ошибки прогнозирования с применением обобщенной модели, о которой было сказано выше.

ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

Краткое рассмотрение применяемых нейросетевых моделей

Рассмотрим *классические полносвязные нейронные сети* (рис. 1). Любая нейронная сеть разделяется структурно на несколько слоев: входной слой, один или несколько внутренних слоев и выходной слой. Каждый слой содержит в себе один или несколько нейронов. Нейрон — это некая функция, которая суммирует все, что приходит на вход по межнейронным связям, полученное значение преобразует функцией активации f и передает на выход нейрона. Количество нейронов во *входном слое* обычно соответствует размерности подаваемого на вход вектора, хотя в некоторых случаях, добавляют еще один нейрон смещения (*bias neuron*). Нейроны смещения не имеют входа. Нейроны входного слоя обычно имеют рефлексивную функцию активации

¹ В оригинальной работе понятие “уникальный набор структурных дескрипторов” описывается одним словом: fingerprints.

² KRR — Kernel Ridge Regression (в русскоязычной литературе: метод гребневой регрессии на ядрах Мерсера).

³ GPR — Gauss Process Regression (в русскоязычной литературе: метод гауссовой регрессии).

⁴ RBF — Radial Basis Function (в русскоязычной литературе: радиальная базисная функция)

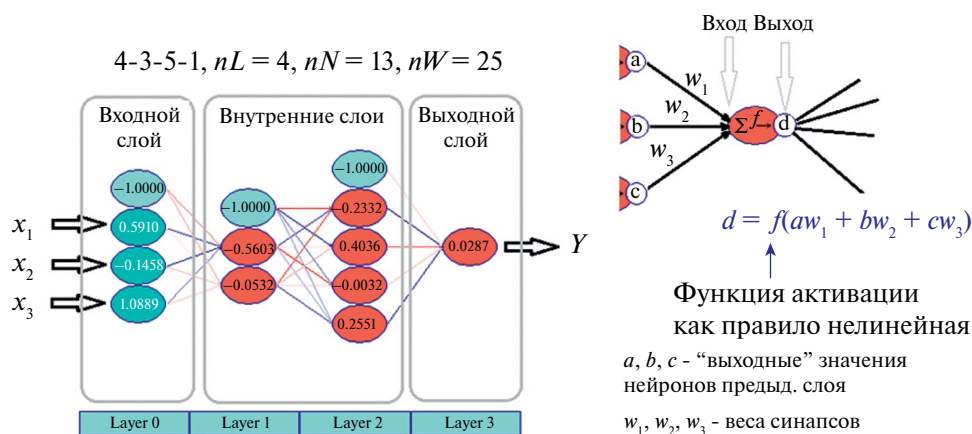


Рис. 1. Схематическое изображение полносвязной классической нейронной сети (MLP) и расчетной схемы нейрона (для простоты изображена нейронная сеть, получающая на вход всего три значения, в реальности их бывает значительно больше, иногда на несколько порядков). MLP – Multilayer Perceptron (в русскоязычной литературе: многослойный перцептрон Розенблатта). Цветные рисунки можно посмотреть в электронной версии.

(т.е. никоим образом не преобразуют входное значение, а сразу передают его на выход нейрона). Каждый нейрон *внутреннего слоя* на вход принимает несколько значений от нейронов предыдущего слоя, которые вначале суммируются, а затем полученное значение подвергается преобразованию с помощью функции активации f (как правило нелинейной) и подается на выход нейрона. В полносвязных сетях каждый выход нейрона связан со входами всех нейронов следующего слоя. Каждая межнейронная связь (синапс) имеет вес w , на который умножается значение с выхода нейрона, перед поступлением на вход нейрона следующего слоя. Проще говоря, нейронная сеть попеременно чередует линейное и нелинейное преобразование. К настоящему времени описано использование более 20 различных нелинейных функций активации нейронов [39], однако наиболее часто применяют либо логистическую функцию активации (она удобна в теоретических исследованиях, но слишком медленно рассчитывается для использования в реальных задачах), либо различные модификации функции ReLU [40].

Все описанное выше относится только к внутренней структуре нейронной сети. Классическая нейронная сеть – в нашем случае это нелинейная функция, так как на выходе у нее будет только одно числовое значение. Перед описанием остаточных нейронных сетей стоит рассмотреть, что собой представляет в математическом плане нейронная сеть, если все функции активации заменить на рефлексивные (см. описание к рис. 4). В этом нам поможет рис. 2. Вывод о линейности полученной модели не удивителен, потому что из математики давно известно, что суперпозиция линейных преобразований – это линейное преобразование.

Рассмотрев внутреннюю структуру классической нейронной сети (многослойного перцептрона) и возможность сведения перцептрона к линейной функции многих переменных путем замены функции активации на рефлексивную, мы вплотную подошли к понятию остаточных нейронных сетей.

Понятие остаточных нейронных сетей обязано своим появлением современной сверточной архитектуре ResNet, появившейся в 2015 году благодаря работе К. Не с сотрудниками [41]. Основная идея этой архитектуры заключалась в добавлении к каждому слою нейронов обходного пути, иными словами, при необходимости слой нейронов просто передает значения своих входов на выходы (по сути, используя рефлексивную функцию активации $y(x) = x$). Эта модификация позволила успешно обучать очень глубокие сети размером до 1000 слоев. До появления ResNet многие исследователи считали обучение нейронных сетей размером более 100 слоев практически невозможным из-за так называемой “проблемы исчезающих градиентов” [39]. В настоящее время почти все современные архитектуры нейронных сетей в том или ином виде реализуют идею обходного пути, однако обычно это довольно глубокие сверточные сети. Мы же попытались реализовать эту идею на небольшой классической полносвязной нейронной сети прямого распространения (многослойном перцептроне Розенблатта [42], модифицированном в соответствии современными представлениями).

По структуре предлагаемая нами остаточная сеть тождественна классической полносвязной нейронной сети, однако каждый внутренний слой должен содержать дополнительно по одному

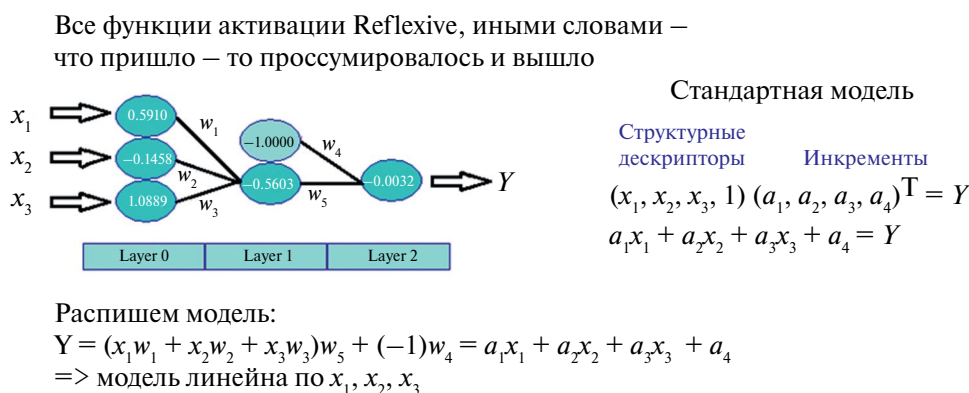


Рис. 2. Математическая тождественность линейной модели и нейронной сети с рефлексивными функциями активации (для простоты рассматривается нейронная сеть только с тремя входными значениями).

нейрону с рефлексивной функцией активации (см. описание к рис. 4), и функции активации входного и выходного слоев должны быть рефлексивны. В математическом виде такую модель можно представить, как

$$Y = \alpha F_{lin}(x_1, \dots, x_n) + (1 - \alpha) F_{nonlin}(x_1, \dots, x_n), \quad (1)$$

где $F_{lin}(x_1, \dots, x_n)$ – линейная функция, $F_{nonlin}(x_1, \dots, x_n)$ – нелинейная функция, $\alpha \in [0, 1]$, причем, если линейная модель функционирует удовлетворительно, то следует ожидать того, что значение α будет близко к единице, иными словами, вклад нелинейной составляющей будет мал.

Описание обобщенной линейной модели

Построение *обобщенной линейной модели* проводили с использованием базы данных ИНХС “Газоразделительные параметры стеклообразных полимеров” [43]. Данные по экспериментальным логарифмам коэффициентов проницаемости газов при 35 °С для каждой из 3052 систем полимер–газ были дополнены 49 структурными дескрипторами полимера, полученными по алгоритмам ранее использовавшимся в модели атомных вкладов [25] с помощью программы RIADA и двумя наиболее подходящими, по нашему мнению, свойствами газа. В качестве свойств газа использовали эффективный параметр потенциала Леннард-Джонса $(\epsilon/k)_{eff}$ и квадрат кинетического диаметра газа d_{eff}^2 , приведенные в работе [44]. Итак, для построения модели составляли переопределенную линейную систему уравнений, в каждом уравнении которой левая часть

представляла собой сумму произведений **числовых** значений структурных дескрипторов A_1 – A_{49} на соответствующие переменные x_1 – x_{49} для полимера и сумму двух произведений $A_{50}x_{50}$ и $A_{51}x_{51}$, относящихся к системе полимер–газ (см. уравнения (2), (3)). Правая часть каждого уравнения системы представляла собой десятичный логарифм коэффициента проницаемости газа, деленный на квадрат кинетического диаметра газа, $(\log_{10} P_{35})/d_{eff}^2$. Решив эту систему уравнений, получим одновременно и параметры модели x_1 – x_{49} для структурных дескрипторов, и оценку констант Тейлора–Мирса – $(K_1 + K_3)$ и K_4 (которые также являются параметрами модели).

$$\frac{\log_{10} P_{35}}{d_{eff}^2} = \sum_{i=1}^{51} x_i A_i, \quad (2)$$

$$\frac{\log_{10} P_{35}}{d_{eff}^2} = \left(\sum_{i=1}^{49} x_i A_i \right) + K_4 \frac{(\epsilon/k)_{eff}}{d_{eff}^2} + \frac{(K_1 + K_3)}{d_{eff}^2}, \quad (3)$$

где $[A_1, \dots, A_{49}]$ – структурные дескрипторы; $A_{50} = \frac{(\epsilon/k)_{eff}}{d_{eff}^2}$; $A_{51} = \frac{1}{d_{eff}^2}$; $[x_1, \dots, x_{51}]$ – переменные ($x_{50} \equiv K_4$, $x_{51} \equiv K_1 + K_3$).

На этапе построения модели x_i – это переменные. После построения модели полученные фиксированные числовые значения будем называть параметрами модели

Расчеты параметров линейной модели (x_1 – x_{51}) проводили с помощью программы RIADA методом многомерной линейной регрессии на основе сингулярного разложения (SVD¹) [45].

¹ SVD – Singular Value Decomposition (в русскоязычной литературе: разложение по сингулярным числам).

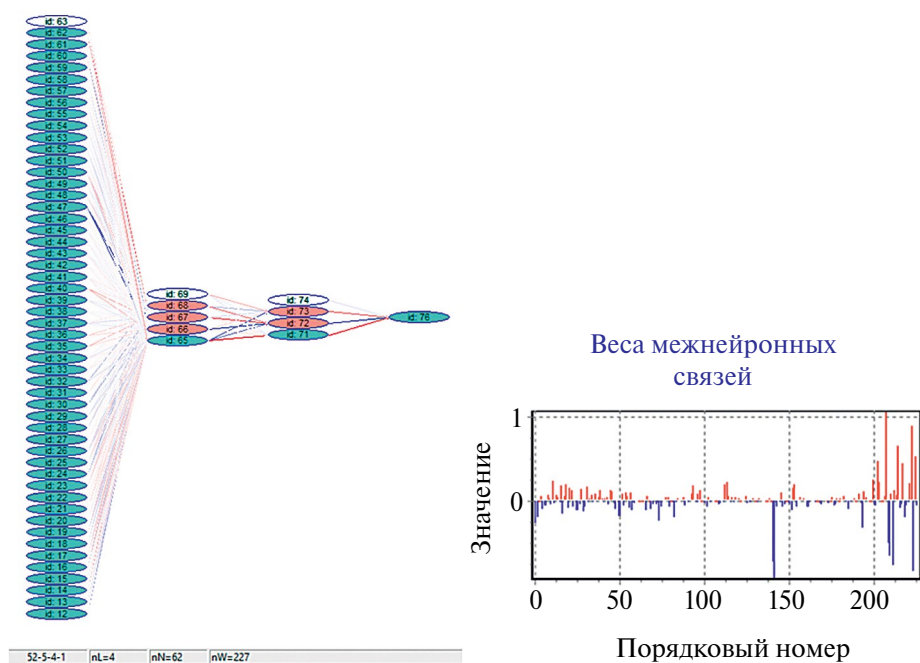


Рис. 3. Структура нейросети (а) и графическое представление весов межнейронных связей (б) после обучения нейросетевой модели, в которой $K_2 = \sum x_i A_i$.

Описание нейросетевой модели

Построение нейросетевой модели проводили с использованием базы данных ИНХС "Газоразделительные параметры стеклообразных полимеров" [43]. Обучающий, валидационный и тестовый наборы данных получены разбиением полного набора данных (3052 системы полимер–газ) с использованием дивергенции Дженсена–Шенона (Jensen–Shannon divergence) [46] так, чтобы функции плотности распределения значений логарифмов коэффициентов проницаемости газов для трех полученных подмножеств были как можно более близкими, а информационная энтропия при этом была как можно выше. Обучающий набор содержал данные для 1534 систем полимер–газ, валидационный — для 901 системы, тестовый — для 617 систем. Нейросетевая модель структурно представляла собой *остаточную трехслойную полностью связную нейронную сеть прямого распространения* (52-5-4-1), содержащую 62 нейрона и 227 весов (рис. 3).

Параметры этой нейросетевой модели приведены на рис. 3б. Видно, что параметры находятся в приемлемом диапазоне (от -1.0 до $+1.0$), т.е. модель способна к активному дальнейшему обучению, и при необходимости может быть дообучена. С другой стороны рис. 3б показывает, что модель близка к сбалансированной, так как количество положительных параметров примерно равно количеству отрицательных, и отсутствуют отдельные резко

выделяющиеся (на несколько единиц) параметры, т.е. не существует явно доминирующих нейронов. Следовательно, количество незначимых для модели параметров является малым.

Внутренние слои сети, помимо нейронов с функцией активации Leaky Rectified Linear Unit (LReLU) [40] (изображены на рис. 3а в виде эллипсов красного цвета), содержали по одному нейрону с *рефлексивной* функцией активации (эллипсы с бирюзовым фоном). Графики функции активации LReLU и рефлексивной функции активации приведены на рис. 4а и 4б соответственно. Входной и первый внутренний слой содержали по одному нейрону смещения (*bias neurons*), изображенных в виде эллипсов с белым фоном. Обучение (около 10000 эпох) проводили методом оптимизации Adam [47]. Для ускорения обучения начальное приближение весов, связанных с линейной частью модели (нейроны 65, 71), взяли из предварительно построенной линейной модели, остальные веса были получены по методу Xavier Glorot [48, 49] в диапазоне от -0.003 до $+0.003$ (т.е. взяты близкими к нулю, но получены современным способом инициализации начального приближения).

Подготовка данных

Входной вектор для каждой системы полимер–газ содержал в себе структурные дескрипторы полимера, полученные с использованием

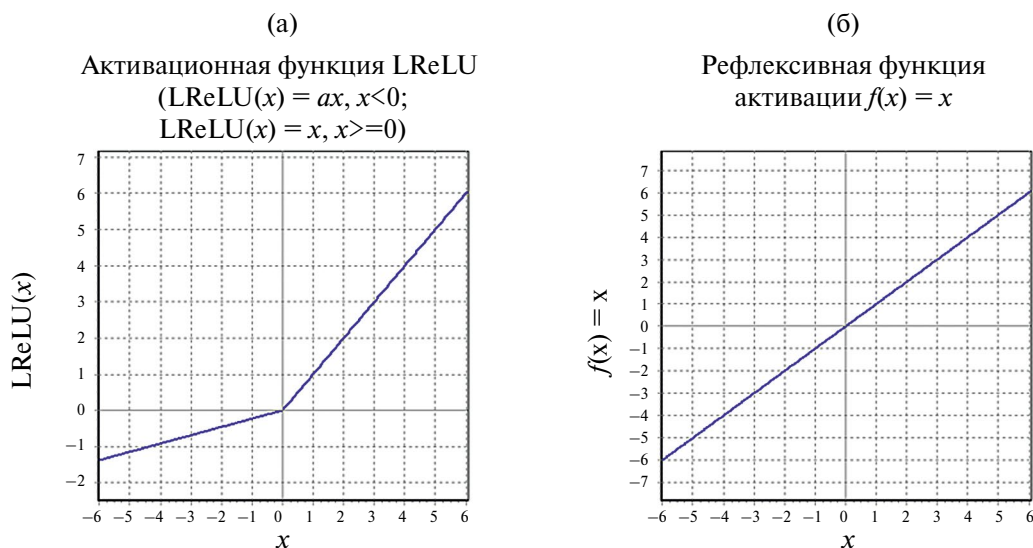


Рис. 4. Графическое представление использованных функций активации нейронов LReLU, при $a = 3/8$ (в русскоязычной литературе часто называется, как ReLU с утечкой) (а) и рефлексивная функция активации $f(x) = x$ (б).

программы RIADA (49 шт.) для полимера и значения, $(\varepsilon/k)_{\text{eff}}/d_{\text{eff}}^2$ и $1/d_{\text{eff}}^2$ для газа (уравнения (2), (3)). Каждую координату входного вектора стандартизовали по всему набору данных. Экспериментальные свойства ($\lg P_{35}$, десятичный логарифм коэффициента проницаемости газа при 35°C) центрировали и шкалировали на диапазон $[-0.85...+0.85]$. При прогнозировании свойства переход к исходному диапазону осуществляли обратным преобразованием.

Обучение нейронной сети

Обучение нейросетевой модели (около 10000 эпох) проводили методом оптимизации Adam [47]. В связи с небольшим объемом исходных данных по сравнению с современными нейросетями обучение осуществляли в многопоточном режиме (48 потоков) на процессоре AMD Ryzen 9 5900X без использования разбиения обучающей выборки на *минибатчи*¹, и оно заняло не более 40 мин. Для предотвращения сильного отклонения отдельных значений от требуемых, **на начальном этапе обучения** использовали метод штрафных функций (penalty function method) в виде использующем функцию с *ограниченным сверху ε -нечувствительным штрафом*: $\text{penalty}(x) = p_1 \exp(-4(p_2/x)^2)$, где x – отклонение от желаемого значения, $p_1 = 100$, $p_2 = 0.2$.

¹ Минибатч – особое понятие, предполагающее не только наличие мининабора, как части обучающей выборки, но и динамического метода его генерации и времени жизни, как правило, ограниченном одной эпохой обучения.

Исключение предпосылок к переобучению

Для предотвращения переобучения использовали *метод ранней остановки* [39]. *Дроп-аут* [50] не использовали в связи с малым количеством весов (227).

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

В рамках используемого подхода возникает необходимость в теоретическом обосновании выбора свойств газа, использованных для построения *обобщенной линейной модели*.

По модели диффузия–растворение проницаемость пенетранта является результатом двух процессов – диффузии через сплошную среду и его растворения в ней [51]. Первый процесс характеризуется коэффициентом диффузии D , второй – коэффициентом растворимости S . Коэффициент растворимости соответствует угловому коэффициенту изотермы сорбции пенетранта в этой среде [52]. Для итогового процесса при невысоких давлениях (активностях), где оба параметра не зависят от давления, можно записать

$$P = D S, \quad (4)$$

где P – коэффициент проницаемости пенетранта, причем каждый из сомножителей (D и S) для фиксированного полимерного образца зависит от свойств газа.

Температурная зависимость коэффициента диффузии пенетранта зачастую линеаризуется в аррениусовских координатах и описывается как

$$\ln D = \ln D_0 - E_D/RT. \quad (5)$$

Здесь E_D — энергия активации диффузии, D_0 — предэкспоненциальный множитель, характеризующий энтропию активации [53]. Р. Меарс предположил, что E_D определяется энергией, затраченной на образование цилиндрической полости (работа против сил когезии):

$$E_D = 1/4 \times \pi d^2 \lambda CED, \quad (6)$$

где d — диаметр цилиндра, соответствующий диаметру частицы пенетранта, λ — длина цилиндра, соответствующего длине диффузионного скачка пенетранта, CED — плотность энергии когезии [53]. Комбинация уравнений (5) и (6) в итоге приводит к общей зависимости, связывающей коэффициент диффузии и диаметр частицы пенетранта:

$$\lg D = K_1 - K_2 d^2, \quad (7)$$

(K_1 и K_2 — константы для фиксированного образца полимера). Для диаметра молекул пенетранта мы использовали шкалу Теплякова—Мирса, построенную на основе тщательного анализа коэффициентов диффузии молекул газов и паров в полимерах, находящихся в различных релаксационных состояниях [44].

Температурная зависимость коэффициента растворимости также линеаризуется в аррениусовских координатах

$$\ln S = \ln S_0 + \Delta H_S/RT, \quad (8)$$

где угловой коэффициент ΔH_S соответствует энтальпии сорбции пенетранта. Энтальпию сорбции можно представить как тепловой эффект суммы двух процессов — конденсации (ΔH_c) и смешения (ΔH_m) [52]:

$$\Delta H_S = \Delta H_c + \Delta H_m. \quad (9)$$

Как правило, $|\Delta H_c| \gg |\Delta H_m|$, и поэтому энтальпия сорбции в основном определяется энтальпией конденсации. Известно, что энтальпия конденсации (теплота испарения) хорошо коррелирует с такими свойствами неполярных пенетрантов, как температура кипения [11] и поверхностное натяжение [54, 55]. В литературных источниках представлено более 20 различных формул, характеризующих связь энтальпии

испарения с критическими параметрами и температурой кипения [56, 57]. Однако следует подчеркнуть, что использование подобного рода сложных зависимостей направлено на обеспечение высокой точности расчета для широкого круга различных по химической природе соединений. В случае ограниченного набора газов и легких паров — при небольших требованиях к точности — можно показать, что теплота испарения коррелирует также с критической температурой, параметром потенциала Леннарда-Джонса (ϵ/k) и инвариантом потенциала Леннарда-Джонса ($(\epsilon/k)\sigma^3/V_m$) (рис. 5). С учетом этого из уравнений (8) и (9) можно получить приближенные выражения для оценки десятичного логарифма коэффициента растворимости:

$$\lg S = K_3 + K_4(\epsilon/k), \quad (10)$$

$$\lg S = h_3 + h_4(\Delta H_{vap}), \quad (11)$$

$$\lg S = t_3 + t_4(T_c), \quad (12)$$

$$\lg S = g_3 + g_4(\gamma_{BP}), \quad (13)$$

$$\lg S = K_3^\# + K_4^\#(\epsilon/k) \sigma^3/V_m. \quad (14)$$

Здесь K_3 , h_3 , t_3 , g_3 , $K_3^\#$ и K_4 , h_4 , t_4 , g_4 , $K_4^\#$ — константы для фиксированного полимера, ΔH_{vap} — энтальпия испарения пенетранта при температуре кипения, γ_{BP} — поверхностное натяжение жидкой фазы пенетранта при температуре кипения, $[(\epsilon/k)\sigma^3/V_m]$ — инвариант потенциала Леннарда-Джонса.

Комбинирование уравнений (4), (7) и (10) приводит к обобщенной линейной зависимости, которая должна выполняться при постановке эксперимента в варианте один полимерный образец—разные газы:

$$\lg P = (K_1 + K_3) - K_2 d^2 + K_4 (\epsilon/k). \quad (15)$$

Анализ значений K_1 , K_2 , K_3 , K_4 для 30 полимеров трех различных типов, проведенный на основе данных из работы [44], позволяет заключить, что K_1 изменяется примерно на 23%, K_3 на 46%, K_4 на 90% (в 1.9 раза), K_2 на 850% (в 9.5 раз) [59]. Несмотря на то, что оценка диапазонов изменения данных параметров весьма приближенная, наилучшим кандидатом для моделирования методом атомных вкладов [25], по-видимому, является параметр K_2 , так как он в относительном плане обладает наиболее широким диапазоном изменения значений. Математическое описание этой модели

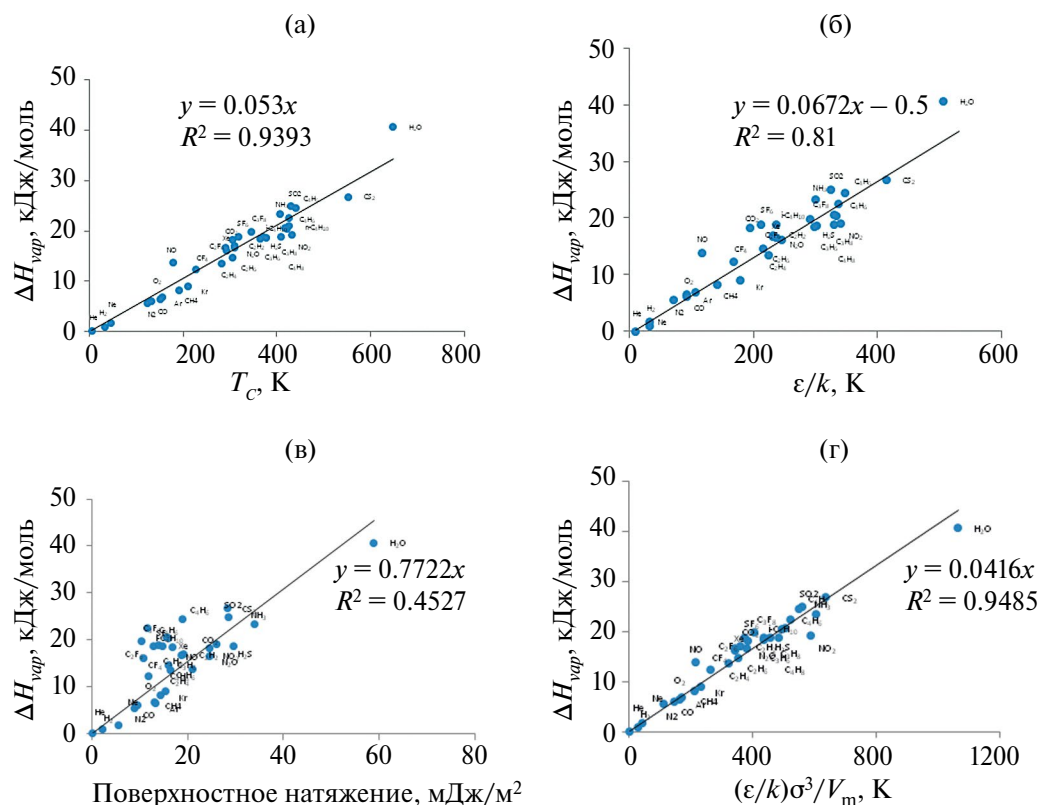


Рис. 5. Корреляция энтальпии испарения ΔH_{vap} с критической температурой T_{cr} (а), параметром потенциала Леннарда-Джонса (ϵ/k), поверхностным натяжением и инвариантом потенциала Леннарда-Джонса (ϵ/k) σ^3/V_m для константных газов и легких паров. Для построения рисунков использованы (ϵ/k) рассчитанные по методике, приведенной в работе [58].

представлено уравнениями (2) и (3), а результаты отражены на рис. 6а.

Помимо данной модели, аналогичным образом была построена модель, в которой параметр K_4 был представлен в виде $\Sigma(x_i A_i)$, причем модель показала крайне низкую сходимость для He, и все 89 точек пришлось отбросить. При этом коэффициент детерминации R^2 составил 0.7445. Для определения причин такого поведения были отдельно построены обобщенные модели для десятичных логарифмов коэффициентов диффузии ($\lg D_{35}$) и десятичных логарифмов коэффициентов растворимости ($\lg S_{35}$). Модели для логарифмов коэффициентов диффузии показали удовлетворительные результаты для всех газов:

$$(8 + \lg D_{35})/(d_{eff}^2) = \Sigma(x_i A_i/d_{eff}^2) - K_2, R^2 = 0.7856, \text{получено } K_2 = 0.501; \quad (16)$$

$$(8 + \lg D_{35})/(d_{eff}^2) = K_1/(d_{eff}^2) + (\Sigma x_i A_i), R^2 = 0.9198, \text{получено } K_1 = 4.454. \quad (17)$$

Одна из моделей для десятичных логарифмов коэффициентов растворимости оказалась приемлемой для всех газов:

$$(2 + \lg S_{35})/(\epsilon/k) = \Sigma(x_i A_i/(\epsilon/k)) + K_4, R^2 = 0.8260, K_4 = 0.0108. \quad (18)$$

Вторая модель, несмотря на хороший коэффициент детерминации, плохо предсказывала коэффициент растворимости гелия:

$$(2 + \lg S_{35})/(\epsilon/k) = K_3/(\epsilon/k) + (\Sigma x_i A_i), R^2 = 0.8853, K_3 = -1.492 \text{ (с He)}. \quad (19)$$

Причины таких отклонений для гелия, с одной стороны, могут быть вызваны большими ошибками определения коэффициента диффузии и, следовательно, коэффициента растворимости гелия, особенно, для средне- и высокопроницаемых полимеров. С другой стороны, причина отклонений может заключаться в некорректном разделении параметров потенциала Леннарда-Джонса,

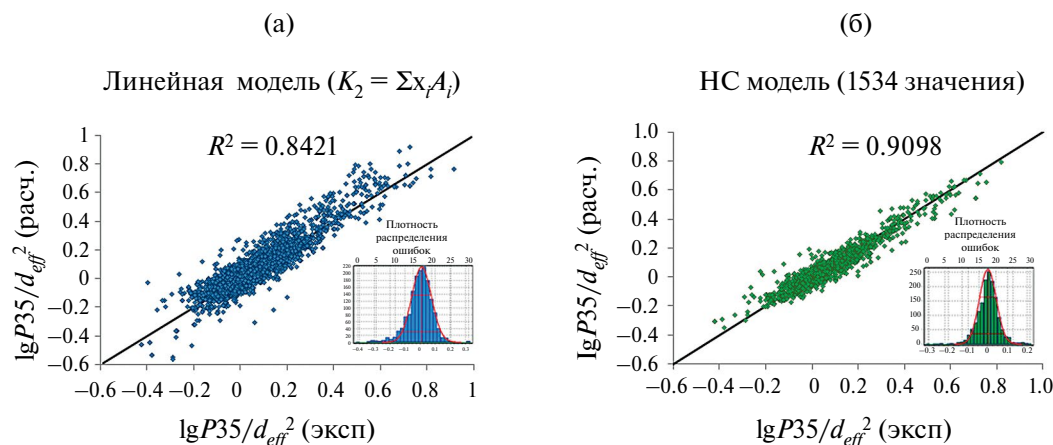


Рис. 6. Сравнение экспериментальных и расчетных значений $\lg P35/d_{eff}^2$ для линейной ($K_2 = \sum x_i A_i$) (а) и нейросетевой (б) моделей. Единицы измерения: $P35$ – Баррер, d_{eff} – ангстремы.

связанном, по мнению авторов книги [60], с самой математической формой потенциала Леннард-Джонса [61], не позволяющей отдельно получить параметры (ϵ/k) и σ :

$$U(r)/k = 4(\epsilon/k) \cdot \left[\frac{\sigma^{12}}{r^{12}} - \frac{\sigma^6}{r^6} \right].$$

Там же рекомендовано использовать оценки (ϵ/k) и σ из критических параметров газов по методике, предложенной в работе [58]:

$$\sigma \left(\frac{P_c}{T_c} \right)^{1/3} = 2.3551 - 0.0874\omega, \quad (21)$$

$$\frac{\epsilon}{kT_c} = 0.7915 + 0.1693\omega, \quad (22)$$

где ω – фактор ацентричности; σ – параметр потенциала Леннард-Джонса, Å; P_c – критическое давление, атм; T_c – критическая температура, К; ϵ – параметр потенциала Леннард-Джонса, эрг; $k = 1.3805 \times 10^{-16}$ эрг/К – постоянная Больцмана.

Использование оценочных значения (ϵ/k) для гелия 3.775 и 30.62 К не привело к существенно улучшению ситуации. Первая оценка была выполнена на основе критической температуры и фактора ацентричности по методике [58], вторая оценка сделана на основе параксора и ψ -фактора в соответствии с формулами приведенными в работе [62]. Таким образом, предположение о некорректном разделении параметров потенциала не нашло своего подтверждения.

В связи с неудачными попытками решить проблему с прогнозом коэффициента растворимости гелия была построена обобщенная модель, в которой K_4 для полимера оценивали на

основе инкрементальной модели атомных вкладов без использования данных по коэффициентам проницаемости гелия. Результаты приведены на рис. 7а.

Далее была построена линейная модель, в которой параметр $(K_1 + K_3)$ был представлен в виде $\sum(x_i A_i)$. Эта линейная модель (рис. 8а) оказалась хуже по сравнению с линейной моделью, представленной на рис. 6а, по коэффициенту детерминации ($R^2 = 0.7627$), поэтому она тоже была исключена из дальнейшего рассмотрения.

Таким образом, из всех моделей, построенных на основе выражения (13), наиболее удачной оказалась линейная модель, в которой K_2 для каждого полимера был представлен суммой атомных вкладов в виде $\sum(x_i A_i)$. Построив модель, на выходе были получены инкременты (x_i) для расчета K_2 и фиксированные значения $(K_1 + K_3) = 3.293$ и $K_4 = 0.01166$, которые, как и инкременты, являются параметрами модели.

Далее была предпринята попытка улучшить линейную модель, путем введения слабой нелинейной составляющей при помощи остаточной нейронной сети. На вход этой нейронной сети подавали те же параметры, что и на вход линейной модели (т.е. структурные дескрипторы $A_1 - A_{49}$ для полимера и два дескриптора $A_{50} = (\epsilon/k)_{eff}/(d_{eff}^2)$ и $A_{51} = 1/(d_{eff}^2)$ для газа). По сравнению с линейной моделью (рис. 6а), нейросетевая модель (рис. 6б), полученная в результате обучения, имеет меньшую погрешность прогноза логарифма коэффициента проницаемости газов и паров при 35 °С на тестовом наборе (в обучении не используется) и более близкое к нормальному распределение остатков.

Аналогичным образом были построены и обучены нейросетевые модели для линейной модели,

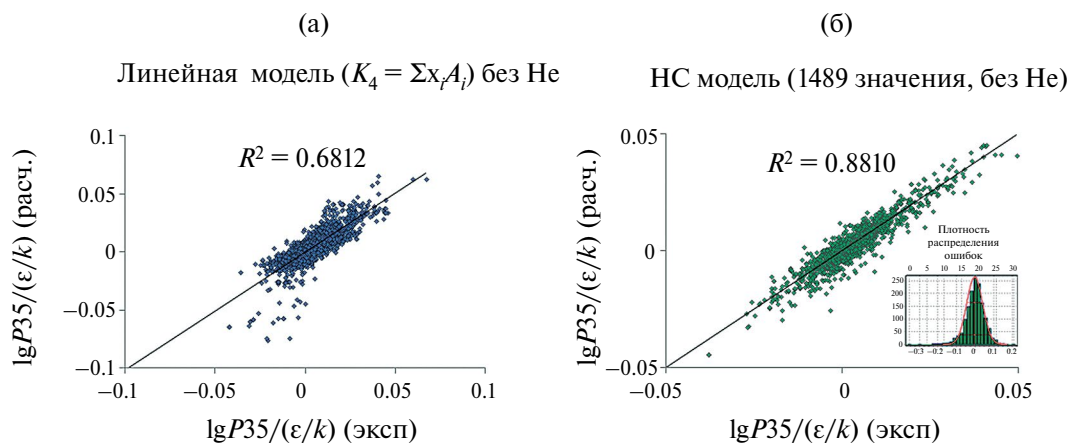


Рис. 7. Сравнение экспериментальных и расчетных значений $\lg P_{35}/(\varepsilon/k)$ для линейной ($K_4 = \sum x_i A_i$) (а) и нейросетевой (б) моделей без He. Единицы измерения: P_{35} – Баррер, (ε/k) – градусы Кельвина

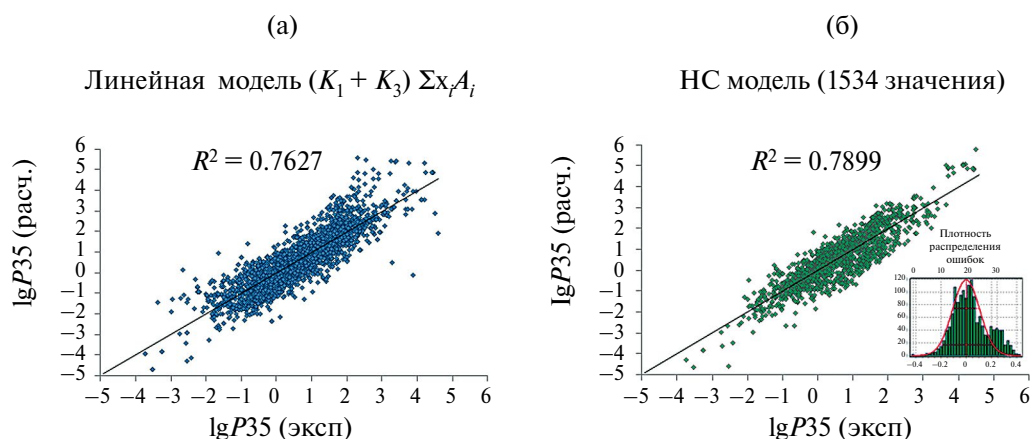


Рис. 8. Сравнение экспериментальных и расчетных значений $\lg P_{35}$ [Баррер] для линейной (а) и нейросетевой (б) моделей.

в которой параметры K_2 (рис. 7б) и $(K_1 + K_3)$ (рис. 8б) были представлены в виде $\sum(x_i A_i)$. Обе модели также позволили улучшить результаты линейной модели (рис. 7а и 8а), однако следует подчеркнуть, что улучшение оказалось не таким значительным, как на рис. 6б. Обучение потребовало более 500000 эпох, несмотря на это, и линейная, и нейросетевая модели оказались хуже модели, представленной на рис. 6б.

Для линейной модели, в которой параметры K_2 были представлены в виде $\sum(x_i A_i)$ и усовершенствованы нейросетевой моделью, было показано, что коэффициенты детерминации в случае конкретных газов (не менее 50 точек для каждой выборки (исходная, обучающая, тестовая, валидационная) имеют удовлетворительные значения ($R^2 \geq 0.80$)).

ЗАКЛЮЧЕНИЕ

Важным результатом данной работы является создание *обобщенной линейной модели* для прогнозирования десятичного логарифма коэффициента проницаемости газов и паров. Если ранее для расчета коэффициента проницаемости каждого газа использовали свои параметры линейной модели, то по итогам текущей работы в *обобщенной линейной модели* параметры x_1 – x_{49} могут применяться к любой системе полимер–газ, в том числе для прогнозирования коэффициента проницаемости неизученного газа, для которого известны параметры потенциала Леннард-Джонса и кинетический диаметр газа (или их оценки). Таким образом, *обобщенная линейная модель на вход получает простейшие структурные дескрипторы для полимера и легкодоступные свойства газа, а на выходе выдает десятичный логарифм коэффициента*

проницаемости газа при 35 °С. Также показана возможность улучшения обобщенных линейных моделей путем введения слабой нелинейной составляющей, с помощью малых остаточных нейронных сетей. При использовании нейронных сетей для улучшения наиболее точной обобщенной модели (с $K_2 = \sum x_i A_i$) коэффициент детерминации возрастает с $R^2 = 0.84$ до $R^2 = 0.91$.

Работа выполнена в рамках Госзадания ИНХС РАН.

СПИСОК ЛИТЕРАТУРЫ

1. Татевский В.М., Яровой С.С., Бендерский В.А. Закономерности и методы расчета физико-химических свойств парафиновых углеводородов : методический материал М.: Гостоптехиздат, 1960.
2. Benson S.W., Buss J.H. // J. Chem. Phys. 1958. V. 29. № 3. P. 546.
3. Benson S.W., Cruickshank F.R., Golden D.M., Haugen G.R., O'Neal H.E. // Chem. Rev. 1969. V. 69. № 3. P. 279.
4. Verevkin S.P. // J. Chem. Eng. Data. 2002. V. 47. № 5. P. 1071.
5. Dorofeeva O.V., Yungman V.S., Saks P. // J. Phys. Chem. A. 2001. V. 105. № 27. P. 6621.
6. Dorofeeva O.V., Ryzhova O.N., Moiseeva N.F. // Russ. J. Phys. Chem. A. 2008. V. 82. № 6. P. 933.
7. Robertson R.E. // Macromolecules. 1969. V. 2. № 3. P. 250.
8. Simha R., Hadden S.T. // J. Chem. Phys. 1956. V. 25. № 4. P. 702.
9. Starkweather H.W., Boyd R.H. // J. Phys. Chem. 1960. V. 64. № 4. P. 410.
10. Слонимский Г.Л., Аскадский А.А., Кутайгородский А.И. // Высокомолек. соед. А. 1970. Т. 12. № 3. С. 494.
11. Van Krevelen D.W., Te Nijenhuis K. Properties of Polymers Their Correlation with Chemical Structure ; their Numerical Estimation and Prediction from Additive Group Contributions. Amsterdam: Elsevier Sci. Technol. Books, 1990.
12. Bicerano J. Prediction of Polymer Properties. New York: Marcel Dekker, 2002.
13. Askadskii A.A., Matveev Yu.I. Chemical Structure and Physical Properties of Polymers. Leningrad: Chemistry, 1983.
14. Askadskij A.A. Computational Materials Science of Polymers. Cambridge: Cambridge Int. Sci. Publ, 2003.
15. Salame M., Steingiser S. // Polymer-Plastics Technol. Eng. 1977. V. 8. № 2. P. 155.
16. Salame M. // Polym. Eng. Sci. 1986. V. 26. № 22. P. 1543.
17. Salame M. // J. Polym. Sci., Polym. Symp. 1973. V. 41. № 1. P. 1.
18. Ямпольский Ю.П., Платэ Н.А. // Высокомолек. соед. А. 1994. Т. 36, № 11. С. 1894.
19. Robeson L. // J. Membr. Sci. 1997. V. 132. № 1. P. 33.
20. Laciak D.V., Robeson L.M., Smith C.D. // Polymer Membranes for Gas and Vapor Separation: Chemistry and Materials Science / Ed. by B.D. Freeman, I. Pinnau. Washington: Am. Chem. Soc., 1999. V. 733.
21. Park J.Y., Paul D.R. // J. Membr. Sci. 1997. V. 125. № 1. P. 23.
22. Yampolskii Yu., Shishatskii S., Alentiev A., Loza K. // J. Membr. Sci. 1998. V. 149. № 2. P. 203.
23. Alentiev A.Yu., Loza K.A., Yampolskii Yu.P. // J. Membr. Sci. 2000. V. 167. № 1. P. 91.
24. Рыжих В.Е. Дис. ... канд. хим. наук. М.: ИНХС РАН, 2014.
25. Ryzhikh V., Tsarev D., Alentiev A., Yampolskii Yu. // J. Membr. Sci. 2015. V. 487. P. 189.
26. Goubko M., Miloserdov O., Yampolskii Yu., Alentiev A., Ryzhikh V. // J. Polym. Sci., Polym. Phys. 2017. V. 55. № 3. P. 228.
27. Alentiev A., Chirkov S., Nikiforov R., Buzin M., Miloserdov O., Ryzhikh V., Belov N., Shaposhnikova V., Salazkin S. // Membranes. 2021. V. 11. № 9. P. 677.
28. Velioğlu S., Tantekin-Ersolmaz S.B. // J. Membr. Sci. 2015. V. 480. P. 47.
29. Volgin I.V., Batyr P.A., Matseevich A.V., Dobrovskiy A Yu., Andreeva M.V., Nazarychev V.M., Larin S.V., Goikhman M.Ya., Vizilter Y.V., Askadskii A.A., Lyulin S.V. // ACS Omega. 2022. V. 7. № 48. P. 43678.
30. Huan T.D., Kim C., Chen L., Chandrasekaran A., Batra R., Venkatram S., Kamal D., Lightstone J. P., Gurnani R., Shetty P., Ramprasad M., Laws J., Shelton M., Ramprasad R. // J. Appl. Phys. 2020. V. 128. № 17. P. 171104.
31. Kim C., Chandrasekaran A., Huan T.D., Das D., Ramprasad R. // J. Phys. Chem. C. 2018. V. 122. № 31. P. 17575.
32. Barnett J.W., Bilchak C.R., Wang Y., Benicewicz B.C., Murdock L.A., Bereau T., Kumar S.K. // Sci. Adv. 2020. V. 6. № 20. P. eaaz4301.
33. Ramprasad R., Batra R., Pilania G., Mannodi-Kanakkithodi A., Kim C. // npj Comput Mater. 2017. V. 3. № 1. P. 54.
34. Tao L., He J., Arbaugh T., McCutcheon J.R., Li Y. // J. Membr. Sci. 2023. V. 665. P. 121131.

35. *Reiser P., Neubert M., Eberhard A., Torresi L., Zhou C., Shao C., Metni H., Van Hoesel C., Schopmans H., Sommer T., Friederich P.* // Commun Mater. 2022. V. 3. № 1. P. 93.
36. *Teplyakov V., Meares P.* // Gas Separ. Purif. 1990. V. 4. № 2. P. 66.
37. *Тепляков В.В., Дургарьян С.П.* // Высокомолек. соед. А. 1984. V. 24. № 7. P. 1498.
38. *Тепляков В.В.* // Журн. Всесоюз. хим. о-ва им. Д.И. Менделеева. 1987. Т. 22. № 6. С. 693.
39. *Николенко С.И., Кадурич А.А., Архангельская Е.О.* Глубокое обучение. Погружение в мир нейронных сетей / Под ред. Н. Гринчик. СПб: Питер, 2020.
40. *Xu B., Wang N., Chen T., Li M.* Empirical Evaluation of Rectified Activations in Convolutional Network. arXiv, 2015.
41. *He K., Zhang X., Ren S., Sun J.* Deep Residual Learning for Image Recognition. arXiv, 2015.
42. *Rosenblatt F.* // Psychol. Rev. 1958. V. 65. № 6. P. 386.
43. ИНХС РАН. База данных "Газоразделительные параметры стеклообразных полимеров". Информрегистр РФ, 1998, № 3585.
44. *Teplyakov V., Meares P.* // Gas Separ. Purif. 1990. V. 4. № 2. P. 66.
45. *Уоткинс Д.С.* Основы матричных вычислений. М.: Лаборатория знаний, 2017.
46. *Lin J.* // IEEE Trans. Inform. Theory. 1991. V. 37. № 1. P. 145.
47. Kingma D.P., Ba J. // arXiv:1412.6980. 2014.
48. *Glorot X., Bengio Y.* // 13th Int. Conf. on Artificial Intelligence and Statistics (AISTATS-10). 2010. P. 249.
49. *Glorot X., Bordes A., Bengio Y.* 14th Int. Conf. on Artificial Intelligence and Statistics (AISTATS-11). 2011. P. 315.
50. *Srivastava N., Hinton G., Krizhevsky A., Sutskever I., Salakhutdinov R.* // J. Machine Learning Res. 2014. V. 15(1). P. 1929.
51. Polymeric Gas Separation Membranes / Ed. by D.R. Paul, Y.P. Yampolskii. Boca Raton: CRC Press, 1994.
52. *Yampolskii Y., Paterson R.* Solubility of Gases in Polymers / Ed. by G.T. Hefter, R.P.T. Tomkins. Chichester: Wiley, 2004. P. 151.
53. *Meares P.* // J. Am. Chem. Soc. 1954. V. 76. № 13. P. 3415.
54. *Keeney M., Heicklen J.* // J. Inorg. Nuclear Chem. 1979. V. 41. № 12. P. 1755.
55. *Alibakhshi A.* // Fluid Phase Equilibria. 2017. V. 432. P. 62.
56. *Svoboda V., Basařová P.* // Fluid Phase Equilibria. 1994. V. 93. P. 167.
57. *Cachadiña I., Mulero A.* // Fluid Phase Equilibria. 2006. V. 240. № 2. P. 173.
58. *Tee L.S., Gotoh S., Stewart W.E.* // Industr. Eng. Chem. Fundamentals. 1966. V. 5. P. 356.
59. *Alentiev A.Y., Ryzhikh V.E., Syrtsova D.A., Belov N.A.* // Russ Chem Rev. 2023. V. 92. № 6. P. RCR5083.
60. *Reid R.C., Prausnitz J.M., Sherwood T.K.* The Properties of Gases and Liquids, Chemical engineering series. New York: McGraw-Hill, 1977.
61. *Lennard-Jones J. E., Devonshire A. F.* // Proc. Roy Soc. London A. 1937. V. 163 (912). P. 53.
62. *Морачевский А.Г., Сладков И.Б.* Физико-химические свойства молекулярных неорганических соединений. Л.: Химия, 1987.