

Федеральное государственное автономное образовательное  
учреждение высшего образования  
«Национальный исследовательский университет ИТМО».

***ДАВЫДОВ Никита Анатольевич***

**Выпускная квалификационная работа**

***Разработка модели искусственного интеллекта для предсказания ширины  
запрещенной зоны слоистых перовскитоподобных оксидов***

Научный руководитель:.

Рецензент:

Санкт-Петербург  
2026

## Оглавление

1	Введение .....	3
2	Обзор литературы .....	6
2.1	Регрессионные модели ИИ для предсказания ширины запрещенной зоны перовскитов .....	6
2.2	Генеративные модели ИИ .....	10
2.3	Регрессионные модели ИИ для предсказания фотокаталитической активности .....	13
3	Теоретическая часть .....	16
4	Обсуждение результатов .....	17
5	Заключение .....	18
6	Благодарности .....	19
7	Список литературы .....	20

## 1 Введение

Поиск оптимальных источников энергии всегда был одной из самых важных проблем для человечества. Сейчас, когда растёт осознание того, что запасы нефти в месторождениях не безграничны, становится актуальной задача поиска нового источника энергии. В этих условиях приобрела популярность концепция водородной энергетики. В качестве топлива предлагается использовать водород, энергию из которого можно извлекать реакции горения с кислородом. В пользу этой идеи говорит и тот факт, что водородная энергетика – это экологически чистая энергетика. Однако одновременно возникает проблема получения водорода в больших количествах. Получение водорода электролизом воды или конверсией метана не могут позволить водороду стать альтернативным источником энергии, так как электроэнергия для электролиза производится, в большинстве своём, за счёт традиционных источников энергии.

Для получения водорода можно использовать солнечную энергию. Преобразователями этого вида энергии в химическую служат фотокатализаторы, которые в большом количестве были синтезированы и изучены на протяжении последних пятидесяти лет и представляют собой специальные вещества, которые делают возможными термодинамически невыгодные реакции (такие как разложение воды), поглощая и передавая реагентам необходимое количество энергии из света.

Одной из существенных характеристик фотокатализаторов является ширина запрещенной зоны: разница в энергии между валентной зоной и зоной проводимости. Именно она определяет возможность протекания той или иной реакции под действием излучения, так как потенциалы соответствующих полуреакций должны лежать внутри диапазона запрещенной зоны, а так же спектр поглощения вещества (<https://doi.org/10.1070/RCR4547>).

Особенное внимание исследователей в отношении фотокаталитического разложения воды привлекли слоистые перовскитоподобные оксиды – кристаллические соединения, в которых последовательно чередуются двумерные блоки (слои) со структурой перовскита ( $\text{CaTiO}_3$ ) со слоями, обладающими иным типом структуры. Их способность к интеркаляции (внедрению молекул и ионов в пространство между слоями со структурой перовскита), в том числе и воды,

обуславливает высокую фотокаталитическую активность по сравнению с ранее известными фотокатализаторами.

Однако, одним из недостатков этого класса соединений является высокое значение ширины запрещенной зоны (3 – 4 эВ), которое приводит к тому, что данные соединения поглощают исключительно в ультрафиолетовой области спектра, в то время, как основная энергия излучения Солнца приходится на более длинноволновую область, делая практическое применение таких фотокатализаторов затруднительным. В связи с этим предпринимались различные попытки понизить ширину запрещенной зоны, путем модификаций оксида. Несмотря на то, что в этом направлении достигнут некоторый прогресс, остается желательным разработать системный подход к такого рода модификациям и, в идеальном случае, предсказывать ширину запрещенной зоны оксида не затрачивая ресурсы на его синтез.

Основным методом теоретического предсказания ширины запрещенной зоны остается квантово-механический расчет на основе методов функционала плотности, которые не лишены недостатков. Помимо довольно большой погрешности (0.3 – 0.5 эВ) ([10.1021/acs.jpcc.1c09594](https://arxiv.org/abs/10.1021/acs.jpcc.1c09594)) такие вычисления требуют значительных вычислительных мощностей и времени, увеличивающихся с усложнением используемых в угоду большей точности функционалов. К тому же, элементарные ячейки слоистых перовскитоподобных оксидов довольно большие и включают несколько десятков атомов, что также приводит к значительной вычислительной нагрузке. Стоит добавить, что проведение квантово-механических расчетов также требует определенной квалификации и навыков, которыми может не обладать рядовой исследователь.

В связи с вышесказанным появляется необходимость в создании инструмента для предсказания ширины запрещенной зоны, который не требовал бы больших вычислительных затрат и был бы легок в использовании, например, модели искусственного интеллекта, обученной на массиве данных определенного класса соединений.

Как было отмечено ранее, погрешность расчетов ширины запрещенной зоны квантово-механическими методами, может достигать 0.3-0.5 эВ, в то время, как значения, полученные экспериментально из спектров поглощения при правильной обработке имеют погрешность порядка 0.03 эВ

(<https://onlinelibrary.wiley.com/doi/epdf/10.1002/pssb.201552007>). Это наблюдение, с одной стороны, делает желательным использование именно экспериментальных, а не рассчитанных значений, в ходе тренировки моделей (в то время как в большинстве работ используются рассчитанные данные), а с другой стороны делает бессмысленным с прикладной точки зрения стремление добиться точности предсказания превышающей 0.03 эВ как погрешность эксперимента.

**Цель работы:**

...

**Задачи работы:**

- ....;

## 2 Обзор литературы

### 2.1 Регрессионные модели ИИ для предсказания ширины запрещенной зоны перовскитов

Область ИИ, предсказывающего свойства материалов, не является новой, а разрабатывается на протяжении длительного времени, в связи с чем накоплен большой объем работ, посвященных этой тематике, в том числе предсказанию ширины запрещенной зоны.

В работе (<https://doi.org/10.1016/j.commatsci.2022.111476>) авторы задались более узкой задачей – классификацией перовскитов по типу электронных переходов в них на прямые и непрямые. Датасет состоял из 1528 строк и включал в себя как оксиды, так и галогениды. В качестве дескрипторов использовался набор, предлагаемый пакетом Matminer. По итогам апробации нескольких моделей () авторы пришли к выводу, что XGBOOST показал наилучший результат (а именно точность в 81%), и что наибольшую роль играют такие дескрипторы как дисперсия расстояния между соседними атомами, минимальное расстояние между соседними атомами, среднее число d-электронов и средняя атомная масса (см. Рисунок 1). Файлы, касающиеся этого исследования, доступны по ссылке (<https://github.com/smarakrath/MI-2021-/tree/main>).

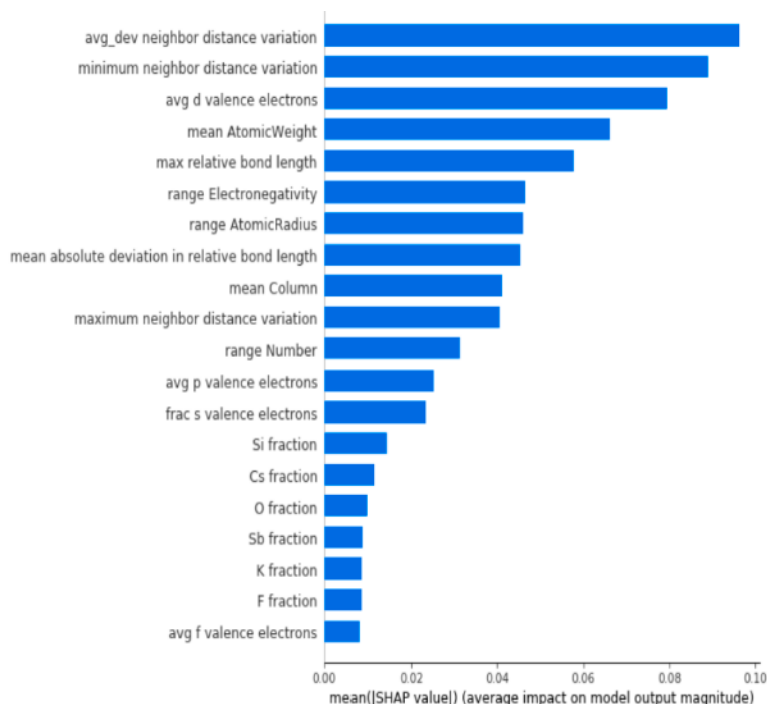


Рисунок 1. Диаграмма значимости дескрипторов из работы (<https://doi.org/10.1016/j.commatsci.2022.111476>)

В работе (<https://doi.org/10.1002/cphc.201800382>) основной акцент был сделан на сравнении предсказаний ИИ и результатов квантово-механических расчетов. Отдельно стоит отметить, что это одна из немногих, где большое внимание уделяется именно слоистым перовскитоподобным оксидам, а не перовскитам в целом. Несмотря на малый объем датасета (56 строк), и относительно простую модель (многослойный перцептрон из 9 нейронов) авторы смогли достичь погрешности предсказания в 3%, используя всего 7 дескрипторов: ионные радиусы и степени окисления атомов в позициях  $ABX_3$  структуры перовскита и число слоев октаэдров, входящих в перовскитный блок. Как видно из статьи, последний из параметров. К сожалению, ссылку на исходный код этого проекта найти не удалось.

По результатам другого исследования (<https://doi.org/10.1021/acs.chemmater.0c02290>) была собрана база данных (<http://www.pdb.nmse-lab.ru/>) из 849 галогенидных перовскитов, содержащая себе кристаллическую структуру, а также ширину запрещенной зоны, рассчитанную DFT и измеренную экспериментально. Затем, на рассчитанных значениях величины ширины запрещенной зоны была обучена модель XGBOOST, использующая комбинацию атомных (атомная масса, положение в периодической системе, электронная конфигурация, ширина запрещенной зоны чистого вещества) и структурных дескрипторов, вычисленных с помощью пакета Matminer (VoronoiFingerprint, CrystalSiteFingerprint, и т.д.). ([https://pubs.acs.org/doi/suppl/10.1021/acs.chemmater.0c02468/suppl\\_file/cm0c02468\\_si\\_002.pdf](https://pubs.acs.org/doi/suppl/10.1021/acs.chemmater.0c02468/suppl_file/cm0c02468_si_002.pdf)). По результатам обучения модель показала средний квадрат отклонения равный 0.14 эВ. Одно из наблюдений, сделанных в ходе этой работы состоит в том, что ширина запрещенной зоны убывает с увеличением количества октаэдров в перовскитном блоке. К сожалению, авторы не приводят ссылку на исходный код, но приводят ссылку на веб-форму модели (<https://eg.scidatasoft.com/>), которая, однако не открывается.

В статье ([10.1039/d1nj01518d](https://doi.org/10.1039/d1nj01518d)) исследователи задались целью предсказывать ширину запрещенной зоны органо-неорганических слоистых перовскитов общей формулы  $A_2BX_4$ . Интересно заметить, что в качестве датасета в данном исследовании выступило подмножество из 136 строк из ранее упомянутой базы данных (<https://doi.org/10.1021/acs.chemmater.0c02290>). Среди 634 изначально рассчитанных

дескрипторов авторы остановили свой выбор на шести из них, которые вместе описывают 93% дисперсии и в дальнейшем были использованы на стадии обучения, а именно: атомные массы галогенид иона и катиона в позиции типа В, logP молекулы, образующей интеркалируемый катион, угол  $\gamma$  элементарной ячейки, MATS4с и GATS4е. Последние два из них относятся к набору дескрипторов PaDEL (<https://onlinelibrary.wiley.com/doi/full/10.1002/jcc.21707>) и связаны с распределением зарядов и электроотрицательностью атомов. Отметим, что 88% вариации в датасете объясняется одним лишь параметром – массой галогенида (I/Br/Cl). Однако, в случае нашей работы это не так релевантно, поскольку мы рассматриваем только оксиды, и присутствие кислорода в позициях типа Х неизменно.

Среди 16 различных моделей ML наименьшую ошибку в 0.09 эВ RMSE показала нейронная сеть с коэффициентом детерминации  $R^2=0.98$ . К сожалению, авторы не приводят ссылку на исходный код, но полностью указывают используемый датасет в приложении (<https://www.rsc.org/suppdata/d1/nj/d1nj01518d/d1nj01518d1.pdf>). Данный датасет состоит из довольно однотипных соединений, где в позиции А находится катион амина, в позиции В катион Pb, Sn, Bi, а в качестве Х, как было сказано ранее, выступает один из трех галогенидов. Это однообразие, по-видимому, и приводит к столь впечатляющей точности предсказания.

Отличный от классического машинного обучения подход был избран в работе (<https://doi.org/10.1016/j.commat.2021.110530>), где была использована модель глубокого обучения MEGNet, основанная на графовых нейронных сетях с attention механизмом. В ходе сбора данных авторы скомбинировали девять различных датасетов, суммарный объем которых после обработки составил более 10 тысяч вхождений. В качестве входа модель MEGNet принимает структуру материала в виде графа, отражающего порядок связей и их свойства ровно как и вершины, отражающие свойства отдельных атомов (см. Рисунок 2). Средняя ошибка на тестовом наборе данных составила 0.28 эВ,  $R^2=0.91$ . В статье приводится ссылка на файлы с моделью (<https://github.com/Pravanop/Perovskite-Prediction>), но создается впечатление, что они не полные, и по большей части описывают подготовку данных перед передачей в модель, в то время, как сам код модели отсутствует, а архив с датасетом поврежден и не открывается.



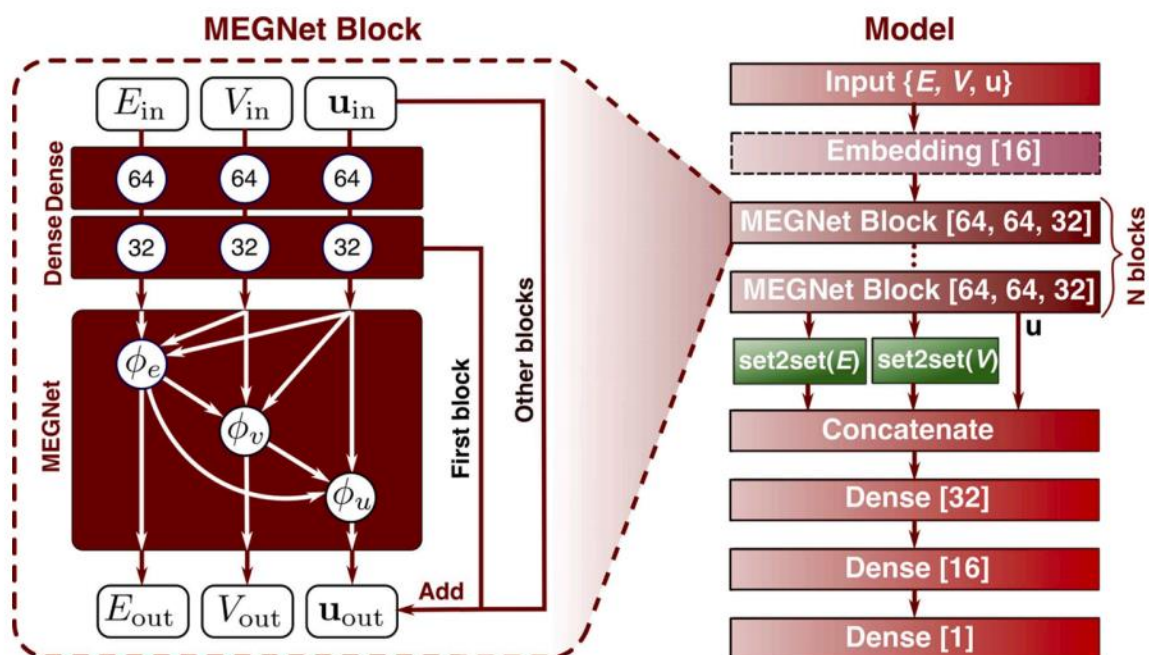


Рисунок 2. Схема архитектуры модели глубокого обучения, используемой в работе (<https://doi.org/10.1016/j.commatsci.2021.110530>)

В работе (<https://doi.org/10.1002/aenm.201901891o>) исследовались следующие модели для предсказания ширины запрещенной зоны галогенидных перовскитов: линейная регрессия, K-ближайших соседей, регрессия на поддерживающих векторах, лес деревьев решений и нейронные сети. На датасете из 69 строк, и one-hot энкодинге состава  $ABX_3$  в качестве дескрипторов наименьшую ошибку (0.06 эВ;  $R^2=0.97$ ) удалось достичь в случае нейронной сети с двумя скрытыми слоями по 8 и 5 нейронов в каждом. В данной работе диапазон составов также не отличается большим разнообразием (A=метиламин, формамидин, Cs; B = Pb, Sn; X = Cl, Br, I) в то время как стехиометрия и кристаллическая структура остаются неизменными, однако в отличие от некоторых других работ в этой использовались исключительно экспериментальные данные. Ссылка на исходный код модели в статье не приводится, но полностью приводится датасет в дополнительных материалах.

В статье (<https://doi.org/10.1016/j.mssp.2023.107427>) на датасете из 199 значений, рассчитанных с помощью функционала HSE06 использовались такие модели как дерево решений, случайный лес деревьев, CatBoost, XGBOOST, LightGBM и многослойный (5 скрытых слоев по 64 нейрона) перцептрон, работающие на 12 дескрипторах: электроотрицательность, ковалентный радиус, первая энергия ионизации и период в периодической системе для каждого из элементов в позиции

ABX<sub>3</sub>. Наилучшими показателями, которых удалось добиться, является  $RMSE = 1.3$  эВ и  $R^2 = 0.70$  для модели CatBoost, что является несколько слабым результатом по сравнению с ранее упомянутыми моделями. В то же время, авторы провели обширный анализ значимости и выявили, что особенно важной в вопросе определения ширины запрещенной зоны является электроотрицательность катиона в позиции типа В. Ссылку на исходный код модели в тексте статьи найти не удалось.

## 2.2 Генеративные модели ИИ для дизайна новых фотокатализаторов

Наряду с задачей предсказания ширины запрещенной зоны на основе известной структуры представляет интерес и обратная задача генерации новых материалов с заданной шириной запрещенной зоны.

В целом, проблеме дизайна материалов с помощью ИИ посвящены несколько обзоров, такие как (<https://doi.org/10.48550/arXiv.2508.03278>). В них выделяются основные подходы к генерации материалов, а именно:

- вариационный автоэнкодер;
- генеративно-сопоставительная нейросеть;
- диффузионные нейросети;
- рекуррентные нейросети;
- трансформеры;
- обучение с подкреплением;

а также способы цифрового представления материалов:

- строковый;
- графовый;
- воксельный;
- физически-обоснованный;
- мультимодальный;

Из этих же обзоров даются ссылки на более узкие работы, посвященные конкретно генерации материалов на основе электронных свойств. Так, в статье (<https://doi.org/10.48550/arXiv.2408.07608>) предлагается модель трансформер MatterGPT (см. Рисунок 3), которая использует строковую репрезентацию кристаллических образцов.

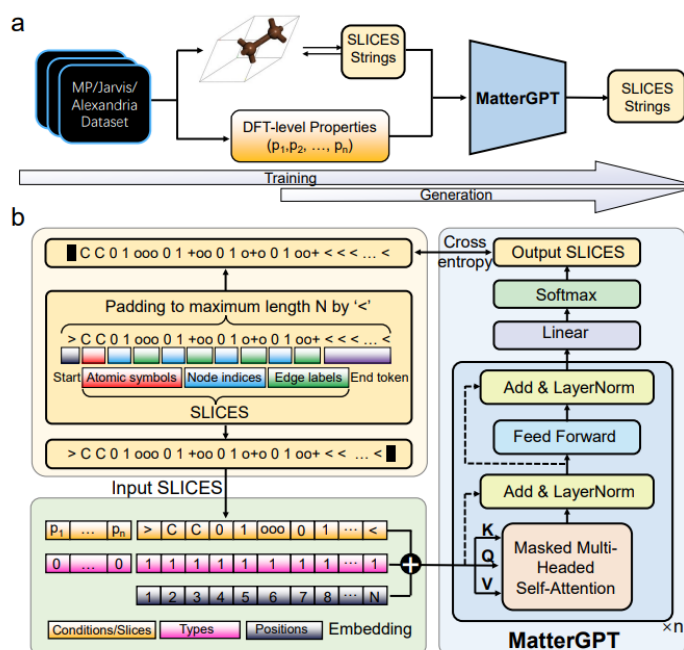


Рисунок 3. Архитектура модели MatterGPT из работы

(<https://doi.org/10.48550/arXiv.2408.07608>)

Данная модель была обучена на более чем 280 тыс. образцов из базы данных «Alexandria». При генерации материалов с заданной шириной запрещенной зоны в диапазоне от 1 до 4 эВ модель показала высокую валидность сгенерированных материалов (выше 90%), но также довольно высокое расхождение (порядка 40%) между заданной шириной запрещенной зоны и шириной запрещенной зоны посчитанной для сгенерированных материалов при помощи DFT. Стоит также отметить, что авторы предоставили ссылку на веб-интерфейс данной модели ([https://huggingface.co/spaces/xiaohang07/MatterGPT\\_CPU](https://huggingface.co/spaces/xiaohang07/MatterGPT_CPU)). По данной ссылке можно также найти файлы с исходным кодом модели.

Еще одной генеративной моделью в области дизайна материалов является CrystalFlow, описанная в статье (<https://www.nature.com/articles/s41467-025-64364-4>). Изначально модель решает задачу предсказания кристаллической структуры для заданного состава, но может также осуществлять генерацию структур при заданном давлении или энергии образования кристалла. Модель была опробована на двух датасетах из Materials Project по 40 тыс. образцов в каждом.

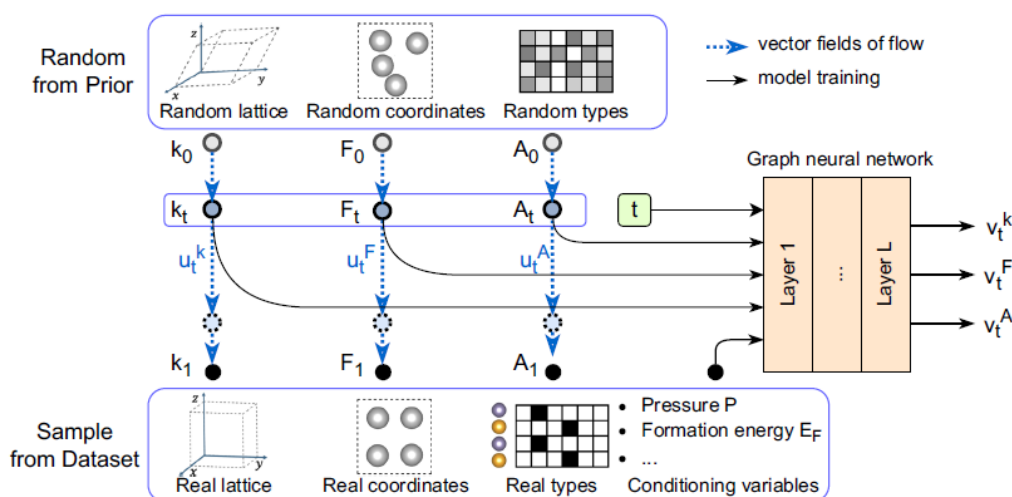


Рисунок 4. Схема архитектуры генеративной модели CrystalFlow, описанной в работе (<https://www.nature.com/articles/s41467-025-64364-4>)

Однако модель показала смешанные результаты. В рамках предсказания кристаллической структуры, а именно 78% совпадения кристаллических структур для кристаллов с менее чем 20 атомами в элементарной ячейке и только лишь 38% для структур с числом атомов в элементарной ячейке достигающим 52.

Что касается результатов условной генерации, то в этом случае ошибка тоже довольно значительная. Так при генерации кристаллов с заданной энергией образования в 1, 2, 3, 4 эВ/атом расхождения с DFT варьировалось от 0.17 до 0.83 эВ/атом. Исходный код модели доступен по ссылке (<https://github.com/ixsluo/CrystalFlow>).

В статье (<https://doi.org/10.48550/arXiv.2406.04713>) предлагается модель FlowMM основанная на использовании Riemannian Flow Matching. Исходный код модели доступен по ссылке (<https://github.com/facebookresearch/flowmm>). Эта модель среди прочего было испытана на тех же двух датасетах, что и предыдущая модель где показала совпадение предсказанной и истинной структур равное 61 и 18% соответственно, что уступает CrystalFlow.

Существует также исследование (<https://doi.org/10.1021/acs.jpcc.1c02438>), в котором генеративно-состязательная нейросеть MATGAN была применена для генерации материалов с заданной шириной запрещенной зоны. Модель была обучена на 66324 соединениях из Materials Project. Представление образцов, подаваемое на вход MATGAN являлось поэлементным one-hot энкодингом. На тестовом наборе данных модель показала следующие метрики  $RMSE = 0.83$  эВ ;  $R^2 = 0.85$  Большая

погрешность объясняется большим разнообразием возможных значений ширины запрещенной зоны в тренировочном датасете и несбалансированностью самого датасета с преобладанием соединений с узкой запрещенной зоной. Авторы также приводят ссылку на исходный код модели (<https://github.com/glard/Active-Generative-Design>).

### **2.3 Регрессионные модели ИИ для предсказания фотокаталитической активности**

Как становится понятно из предыдущих разделов, предсказание ширины запрещенной зоны перовскитов довольно разработанная тема, хотя и не без вызовов для дальнейшего прогресса. Однако, гораздо более привлекательной является предсказание скорости выделения водорода как конечного показателя эффективности фотокатализатора. В отличие от предыдущей задачи эта проблема является существенно менее исследованной в силу своей узкой направленности, и на порядки меньшего количества данных, которые, к тому же, могут быть определены только экспериментальным путем. На это также накладывается фактор отсутствия стандартных условий проведения фотокаталитического эксперимента, из-за чего каждый научный коллектив собирает свои «кустарные» установки, проводит эксперимент при своих условиях, которые, в случае измерения фотокаталитической активности значительно сильнее влияют на измеряемую величину, чем в случае ширины запрещенной зоны, которая, по большому счету, определяется исключительно природой образца. Даже при соблюдении постоянных условий в рамках одной и той же научной группы, одно и то же вещество в форме образцов в двух разных статьях может демонстрировать значительные расхождения в активности, что говорит о непредсказуемости влияния большого количества факторов, о которых можно даже не догадываться. Тем не менее, некоторые работы по этой теме все же встречаются.

Так, в статье (<https://doi.org/10.1016/j.jechem.2021.01.035>) использовался датасет из 77 значений собранных из различных других работ, в котором помимо состава перовскита приводятся такие условия эксперимента как температура, используемая лампа и концентрация жертвенного агента. В качестве дескрипторов выступали 18 величин описывающих атомные свойства входящих в состав катализатора элементов

и условия синтеза, а также экспериментальные условия, кодируемые one-hot энкодингом из шести вариантов. В ходе отбора дескрипторов было выявлено, что наибольшую роль играют методика синтеза и природа катиона, находящегося в позициях типа А. Среди нескольких моделей ИИ наилучшие метрики продемонстрировала модель байесовой нейронной сети (BRANN), а именно  $RMSE = 371$  мкмоль  $H_2/(ч*г)$ ;  $R^2 = 0.97$  на тестовом наборе данных. Авторы указали ссылку на веб-сервис с данной моделью, которая, однако, не работает ([http://materials-data-mining.com/ocpmdm/material\\_api/i0ucuyn3wsd14940](http://materials-data-mining.com/ocpmdm/material_api/i0ucuyn3wsd14940)), а также полностью привели датасет в дополнительных материалах.

В новом и объемном обзоре (<https://doi.org/10.1039/D5NJ00505A>), тематикой которого является в целом применение ИИ к проблемам фотокаталитического разложения воды, также уделяется некоторое внимание работам, нацеленным на предсказание скорости выделения водорода. Так, в этом обзоре дается ссылка на статью, авторы которой собрали и целиком привели в дополнительных материалах датасет из 540 строк со значениями скорости выделения водорода (<https://doi.org/10.1016/j.apcatb.2018.09.104>) в зависимости от таких параметров как элементный состав в формуле  $ABX_3$ , информация о методе синтеза, сингонии, удельной площади поверхности и довольно детальном описании условий эксперимента (загрузка катализатора, концентрация жертвенного агента, тип и мощность источника света). Они также испытали несколько ИИ моделей на этом датасете, лучшая из которых (случайный лес) показала  $RMSE = 1194$  мкмоль  $H_2/(ч*г)$ ;  $R^2 = 0.79$ .

Есть также множество работ, где модели машинного обучения строятся для предсказания фотокаталитической активности конкретного фотокатализатора в зависимости от условий проведения эксперимента (<https://doi.org/10.1016/j.enconman.2023.118007>), (<https://doi.org/10.1016/j.ijhydene.2024.04.173>), (<https://doi.org/10.1039/D3EY00246B>) или выбора допанта (<https://doi.org/10.1002/adma.202106776>), но это не совсем то направление, которое мы выбрали в ходе данной работы, ставящей своей целью установление взаимосвязи «структура-свойство», а не «условия эксперимента-свойство».

В исследовании ([10.1038/s41598-025-10785-6](https://doi.org/10.1038/s41598-025-10785-6)) была применена графовая нейронная сеть для предсказания ширины запрещенной зоны фотокатализаторов на основе композитов  $\text{TiO}_2/\text{ZnO}$  в разном соотношении компонентов. По итогам работы была достигнута точность предсказания равная 0.045 эВ. К сожалению, методика не приводится подробно, а описывается в общих чертах. Касательно использованного датасета указывается только, что он содержит 10000 строк, информацию о составе и структуре образцов, подробное описание архитектуры модели опускается, ровно как не приводится ссылок на код или датасет.

Эту модель также использовали для предсказания эффективности фотокатализатора (хотя и без определения, что под этим понимается) с точностью в 4.3%.

## 2.4 Выводы по результатам обзора литературы

Существует множество регрессионных моделей для предсказания ширины запрещенной зоны различных кристаллических материалов, многие из которых добиваются хорошей точности предсказания. Однако, некоторые исследования фокусируются на узком классе перовскитов, обычно классических со структурой  $\text{ABX}_3$ , в то время, как слоистые перовскитоподобные оксиды показывают большую фотокаталитическую активность, чем классические соединения и предсказание их ширины запрещенной зоны было бы более актуальной проблемой для этой области. Нет гарантии того, что модели, обученные на классических перовскитах, продемонстрируют такую же точность на слоистых материалах в силу различий в структуре и большого разнообразия в составе межслоевого пространства. В то время как в классических перовскитах основной акцент делается на соединения на аммонийные галогениды свинца и олова слоистые перовскиты демонстрируют большее разнообразие как структур так и составов.

Кроме того, значительная часть исследований полагается на расчетные данные, которые, как было сказано выше, страдают от погрешности. В силу этого более корректным выглядит обучение моделей на экспериментальных данных. Так, можно было бы взять датасет из работы (<https://doi.org/10.1016/j.apcatb.2018.09.104>) и совместить с датасетом исключительно по слоистым перовскитоподобным оксидам и обучить модель совместно на них.



Расширение темы работы с ширины запрещенной зоны на предсказание скорости выделения водорода выглядит еще более интересно, поскольку этой проблеме посвящено мало работ и ни одна из них не исследует конкретно слоистые перовскитоподобные оксиды, которые, как было отмечено выше, довольно эффективные фотокатализаторы заслуживающие включения в датасет. К тому же имеющиеся подходы не вдаются в структуру соединений и межатомные взаимодействия, а используют атомные дескрипторы. Можно было бы сделать модель фокусирующуюся на слоистых перовскитоопобных оксидах (или даже просто на перовскитах), но попробовать улучшить точность предсказания скорости выделения водорода учитывая структуру веществ и подробное описание условий эксперимента. К имеющемуся датасету из 540 обычных перовскитов из той же работы (<https://doi.org/10.1016/j.apcatb.2018.09.104>) можно было бы присоединить данные по слоистым перовскитам, которые с большой долей вероятности можно собрать в количестве 100-200 строк. В отличие от ширины запрещенной зоны, которая лежит в пределах 0-5 эВ, скорость выделения водорода может разниться на порядки и в этой задаче больше вызова и места для улучшения существующих решений. В то же время отказываться от решения задачи на предсказание ширины запрещенной зоны, наверное, не стоит, поскольку предсказание скорости выделения водорода может не удался, и тогда придется довольствоваться шириной запрещенной зоны. Поэтому работу можно было бы назвать в духе: «Разработка модели ИИ для предсказания фотокаталитических свойств слоистых перовскитоподобных оксидов (или перовскитов в целом)».

Что касается генеративных моделей, то эта задача еще более интересная, но и гораздо более сложная, чем регрессия. Генерировать валидные кристаллические структуры довольно тяжело, поэтому в условиях недостатка времени декларировать это как тему, наверное, не стоит. Если останется время и удастся добиться каких-то результатов в этом направлении, то это будет приятным бонусом, но пока что я бы остановился на регрессии.

### **3 Теоретическая часть**



## **4 Обсуждение результатов**

## **5 Заключение**

В ходе работы мы пришли к следующим выводам:

- ....;

## **6 Благодарности**

## **7 Список литературы**