**Применение машинного обучения на основе маппинга данных о физических свойствах сплавов Ti и Zr для предсказания упругих свойств**

***Яблонский И.Д.\*1, Бакланов Л.В.1***

*Студент, 2 курса магистратуры*

*1 Магистерская школа Информационных бизнес-систем, Национальный исследовательский технологический университет МИСиС, Москва, Россия*

*E-mail: m2104198@edu.misis.ru*

***Научный руководитель*** *– Сиднов Кирилл Павлович*

Одним из наиболее точных, но в то же время ресурсоемких методов расчета механических свойств кристаллов является метод проекционно-присоединённых плоских волн, но, часто, задачи поиска перспективных материалов требуют множество расчетов. Для ускорения решения таких задач в последнее время используются методы машинного обучения, которые, при наличии достаточной обучающей выборки, способны предложить точность, сопоставимую с расчетными методами [[1,2]](https://paperpile.com/c/6ON5EY/H9ij+200b).

В данной работе предложена система машинного обучения для прогнозирования упругих свойств ОЦК-сплавов Ti и Zr на основе результатов расчетов с низкой точностью. Предсказываемые величины включали 7 независимых упругих констант, модули объемной упругости, сдвига, Юнга и *C’=(C11-C12)/2* для 60 ОЦК-сплавов на основе Ti и Zr.

Модель машинного обучения в рамках метода градиентного бустинга на основе библиотеки “Catboost” [[3]](https://paperpile.com/c/6ON5EY/tlIv) была оптимизирована путем перекрестной проверки по сетке параметров (“Grid Search”). Для преобразования информации о качественном и количественном составе сплавов в векторы признаков была использована библиотека “Matminer” [[4]](https://paperpile.com/c/6ON5EY/akcz). Полученные векторы дополнялись значениями модуля объемной упругости, энергии и объема основного состояния, полученные методом PAW-SQS для 128-атомных ячеек при упрощенных параметрах расчета (1-ая зона бриллюэна описывалась одной гамма точкой). Оценочные затраты на получения этих значений составили 4.3 ядро-часов на состав. Объясняемые величины были получены с помощью точных расчетов, ресурсоемкостью 300 ядро-часов на состав.

Валидация моделей выполнялась в 3-ёх сценариях применения: предсказание характеристик для новых концентрационных сочетаний в системах уже имеющихся в обучающей выборке; предсказание характеристик для систем, которые не были представлены в обучающей выборке; предсказание характеристик для составов, ни один химический элемент из которых не был представлен в составах обучающей выборки.

Средний коэффициент детерминации при прогнозировании упругих характеристик составил 0.64-0.82, средняя абсолютная ошибка 6.5-8 ГПа.

Работа поддержана Российским Научным Фондом, грант № 21-72-10105.

**Литература**

[1] S. Xiong et al. A combined machine learning and density functional theory study of binary Ti-Nb and Ti-Zr alloys: Stability and Young’s modulus, Comput. Mater. Sci. 184 (2020) 109830.

[2] X. Liu et al. Machine Learning Assisted Prediction of Microstructures and Young’s Modulus of Biomedical Multi-Component β-Ti Alloys, Metals . 12 (2022) 796.

[3] [CatBoost - state-of-the-art open-source gradient boosting library with categorical features support, (n.d.).](http://paperpile.com/b/6ON5EY/tlIv) <https://catboost.ai/> [(accessed February 16, 2024).](http://paperpile.com/b/6ON5EY/tlIv)

[4] L. Ward et al. Matminer: An open source toolkit for materials data mining, Comput. Mater. Sci. 152 (2018) 60–69.