

Программа дисциплины

Машинное обучение в моделировании материалов

Программу составил: к.ф.-м.н. Антропов А.С.

Аннотация

Дисциплина посвящена рассмотрению методов машинного обучения (ML) и их применению в физике и науках о материалах. Рассматриваются подходы к построению моделей предсказания свойств материалов, построению машинно-обученных потенциалов межатомного взаимодействия, а также некоторые другие приложения машинного обучения в физических задачах. Особое внимание уделяется использованию ML в многомасштабном моделировании, работе с квантово-механическими данными, а также применению современных нейросетевых архитектур. Практическая часть курса включает работу с базами данных материалов, обучение различных моделей и их интерпретацию.

Цель дисциплины

Целью дисциплины является ознакомление студентов с основами машинного обучения и его приложениями в физике материалов.

Задачи дисциплины

- Ознакомление студентов с основами машинного обучения и его применением в материаловедении.
- Знакомство с различными источниками данных в науках о материалах.
- Освоение методов выбора и генерации дескрипторов для атомных структур.
- Ознакомление с нейросетевыми архитектурами и их применением в задачах материаловедения, в частности с графовыми нейросетями, сверточными нейросетями и трансформерами.
- Практическое обучение работе с библиотеками машинного обучения.

Содержание

- 1) Введение в машинное обучение для наук о материалах. Типовые задачи ML в физике материалов. Процедура обучения моделей: train, test, validation. Источники данных: эксперименты, моделирование, квантово-механические расчёты. Различные масштабы моделей: MLIP, предсказание энергии релаксированной структуры, предсказание свойств материалов. Влияние состава и структуры на свойства материалов.
- 2) Простые ML-модели в материаловедении. Модели состав-свойство: случайный лес. Практика: обучение случайного леса на подвыборке из базы данных проекта Materials Project (MP). Физически интерпретируемые модели. Понятие выпуклой оболочки, энергия образования и абсолютная энергия. Линейная регрессия в предсказании энергий материалов. Практика: обучение линейной регрессии на MP.
- 3) Дескрипторы материалов. Основные понятия: инвариантность, эквивариантность, локальность. Практика: выделение дескрипторов SNAP и GAP для атомов в структурах.
- 4) Введение в нейросетевые методы. Основные архитектуры нейросетей, обратное распространение ошибки, оптимизаторы, планировщики скорости обучения.
- 5) Машинно-обученные потенциалы. Иерархия ML-потенциалов. Обучение моделей на дескрипторах.

- 6) Методы активного обучения: ансамблевый метод, пространство внутренних представлений.
- 7) Графовые нейросети и эквивариантность. Современные универсальные потенциалы.
- 8) Свёрточные нейросети (CNN) в материаловедении. Определение структуры низкомолекулярных соединений по ИК-спектрам.
- 9) Трансформеры для построения SMILES по спектрам молекул.
- 10) ML-модели для предсказания и анализа электронной плотности.