The Materials Project

COS 472

Ben Franklin, Lucas Matheson, Colby Wirth

October 10, 2024

Table of Contents

Overview

What is the Materials Project?

- Multi-national
- Multi-institutional
- Collaborative project aimed at computing the properties of new materials
- Enables expedited scientific research by eliminating costs associated with conducting experiments on compounds

Motivation

Why contribute to the Material Project?

- Test and Implement our new skills in a meaningful manner
- Contribute to a project much greater than ourselves
- Learn more about Chemistry (and Machine Learning!)
- Establish inter-department relationships!

Goals

- Goal 1:
  - Initially implement a ML model that can predict a handful of already known attributes with a high level of precision (greater than 95% ideally)
- Goal 2:
  - Predict unknown attributes for a large subset of compounds from the original dataset

Methods

- Evaluate which attributes are predictable and which are computable
  - We will not predict values that are easily computed.
- Implement a Knowledge Discovery in Database (KDD) pipeline
  - Preprocess data
  - Choose and train ML models
  - Evaluate and tune parameters
- Repeat for Goals 1 and 2
- Analyze results for the final paper

Evaluation Criteria

- Goal 1:
  - Achieve 95% accuracy rate with predicting 'band-gap' (and potentially other) attributes across $\sim$ 50,000 compounds
- Goal 2:
  - Predict 'band gaps' for a large amount of the unknown attributes
- If time permits, repeat Goals 1 and 2 for other attributes.

End

End of Presentation

Questions?