Data Mining Project

Discovery of nano-materials(perovskite) for Green Hydrogen production application using machine learning models

August 14, 2020

Overview

The time required by Nano Materials to enter the Manufacturing sector takes years as a consequence of continuous research and experimentation costs and facilities. With the growing industrial demands, the need for novel materials, with application-specific properties, is increasing rapidly. Considering computational material science techniques, the traditional DFT method is generally used but has a major disadvantage. It requires a large amount of computational power and time. By using Machine learning algorithms approach, efficient prediction of many properties of materials can be done, using only the much easily obtained elemental, structural or compositional information of materials.

Goals

1. **Feature selection for property prediction:** Dataset available to us contains about 70 features. Features specific to the property must be segregated.
2. **Model selection for property prediction:** Both Regression and Classification can be used to predict the stability of materials with various ML algorithms. We have to check the parameters such as accuracy, precision and recall for classification and RMSE and MAE for Regression.

Specifications

Property predicted for the perovskites materials provided in the reference papers is regarding its stability and we are aiming to use the same data for the prediction of various different properties(ex. their ability to generate Hydrogen gas).

Milestones

1. Data Collection

   Data collection of perovskites regarding their property of Hydrogen production from various research papers.

2. Prediction of Materials

   Predicting materials with the specified property.

3. Making feature-wise models

   Specific properties of materials require specific ML models, we aim to make separate algorithms for different material properties.