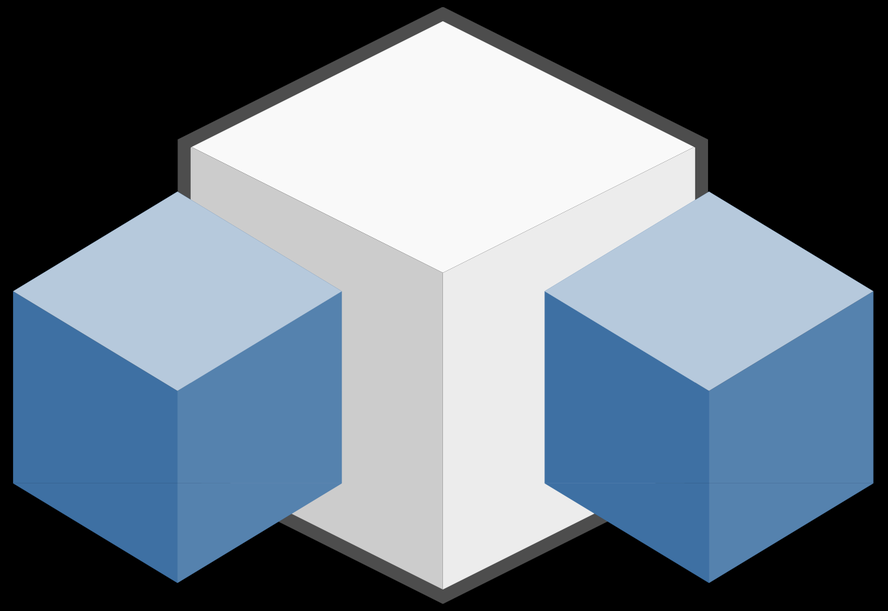
Prediction of Orthorhombic Lattice Constants

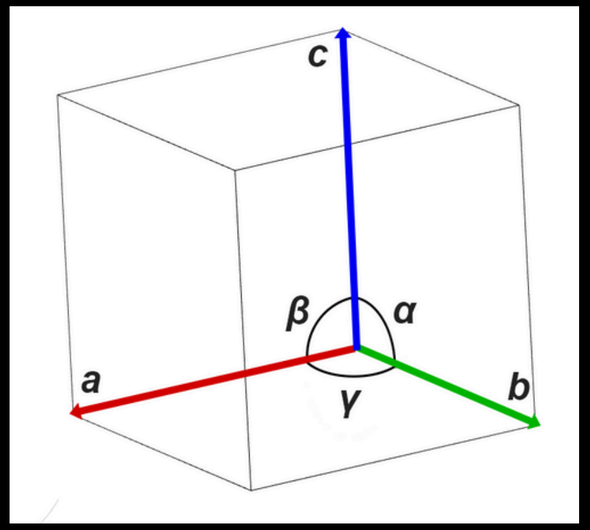


But why?

The structure of crystal materials is denoted by the lattice structure.

The lattice structure is a repeating pattern. Its basic unit is the unit cell.

The unit cell is fully desribed by the six lattice constants.



These are:
a, b, c;
alpha, beta, gamma.

Many properties of the crystal material depend on the **lattice structure**, which depends on the **unit cell**, which depends on the **lattice constants**.

Objectives:

- predict orthorhombic lattice constants
- examine feature importance

Method to the Madness

Thou shalt predict.

Method to the Madness

Dataset: Materials Project

Pre-processing: pandas, numpy

Model: random forest regression (from sklearn)

Post-processing (feature importance): shap

Visualization: matplotlib, shap

Method to the Madness

The dataset is taken from the Materials Project database and modified by Li et al.

It has been featurized using MAGPIE and the matminer library. (insert database shape)

Method to the Madness

The dataset was already clean (NaN values, etc.) so I proceeded to filtering out the orthorhombic crystal materials. (I then perform further cleaning by removing the outliers. More on this later.)

Method to the Madness

I passed the dataset through a function to separate features and targets.
I performed an 80-20 train-test-split for the features and targets.

Method to the Madness

I used sklearn's random forest regressor
class to create three RFR models
with n_estimators=300 and
max_features=300, 200, 100
respectively.

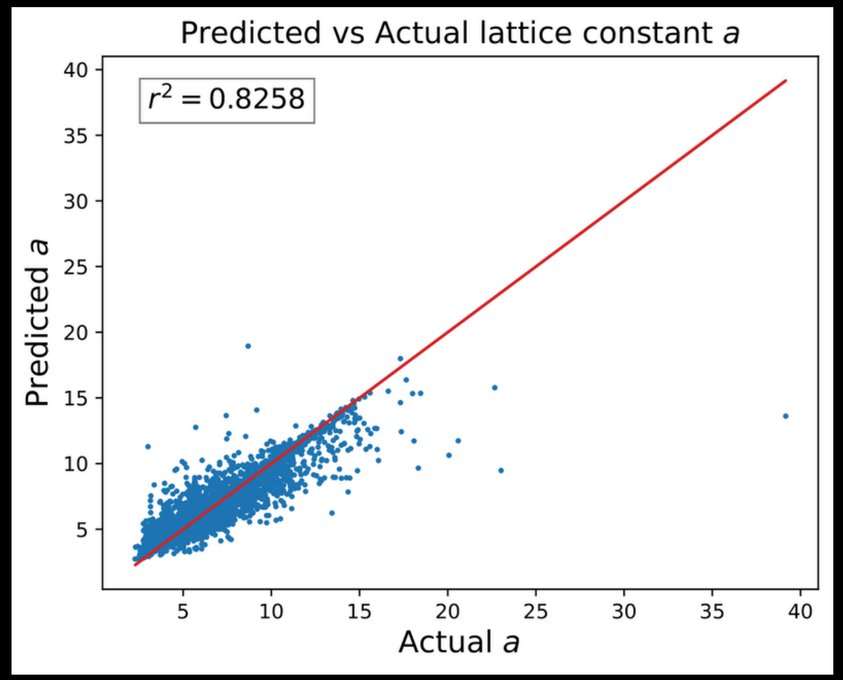
Method to the Madness

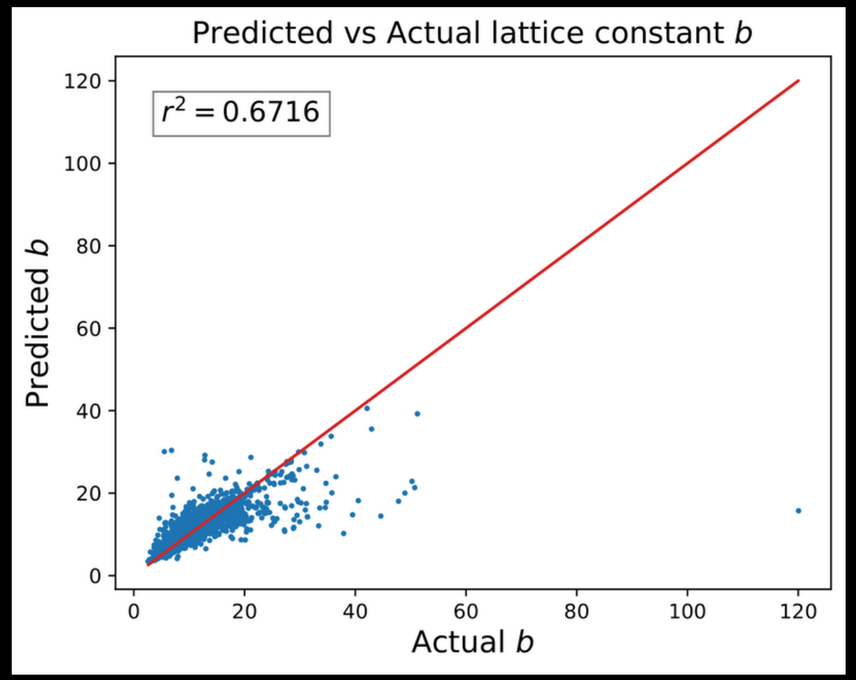
I took the mean of the result of each RFR model
as my final prediction.

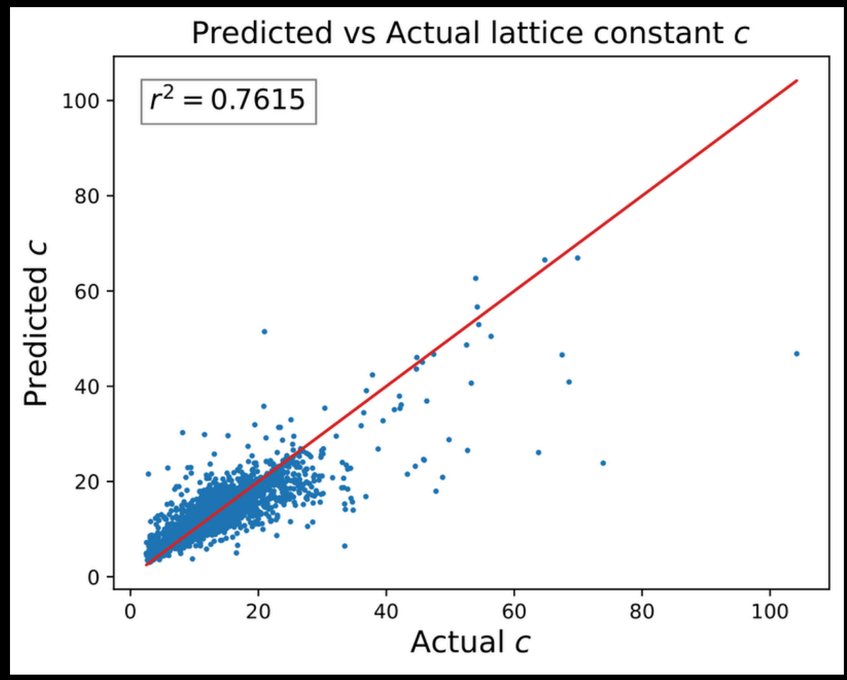
Method to the Madness

I took three metrics: r-squared score, mean absolute error, and mean squared error to gauge the performance of the model.

Fruits of Labor

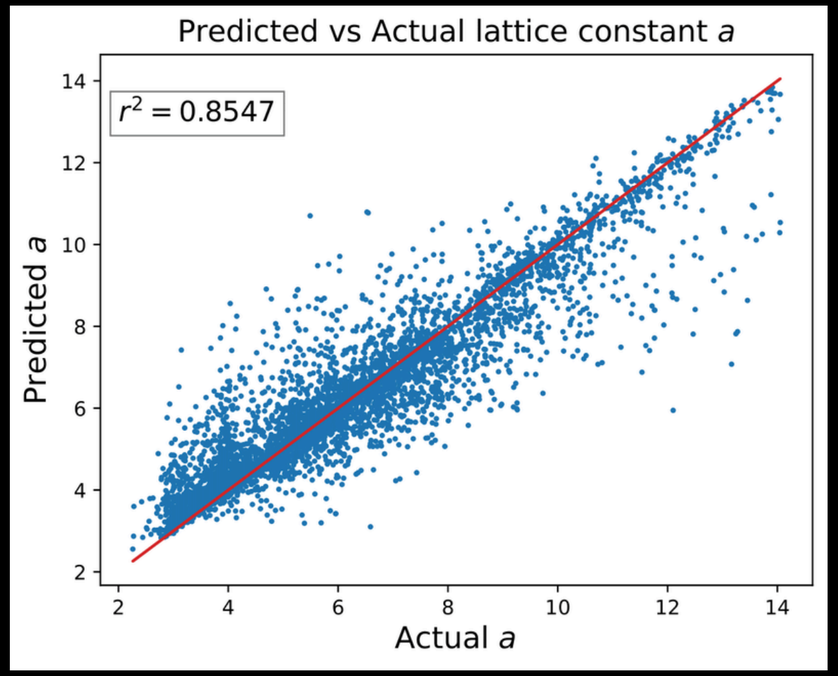


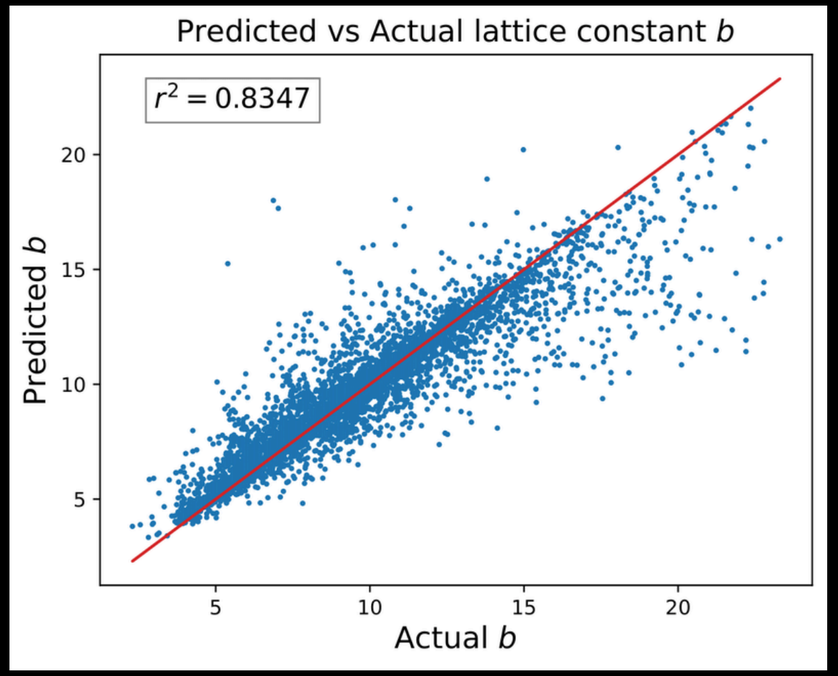


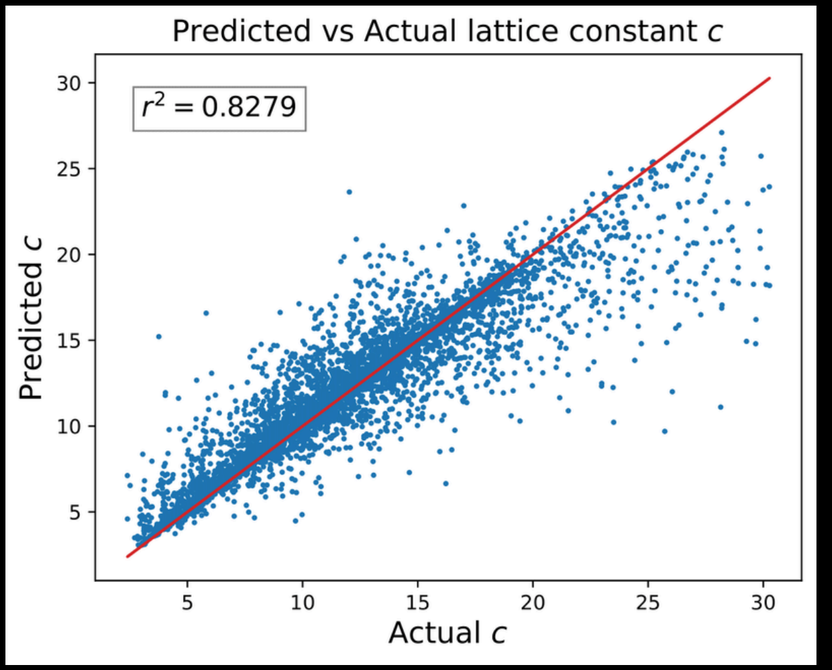


Daffon - Prediction of Orthorhombic Lattice Constants

Fruits of Labor







Daffon - Prediction of Orthorhombic Lattice Constants

What did we learn?

Through random forest regression
we can predict orthorhombic
lattice constants to an acceptable
degree of accuracy.

What did we learn?

By cleaning the outliers in the training data, we improve the performance of our model.

What did we learn?

However, this predictive model
cannot replace DFT ab initio calculations.
It can only supplement the investigation
methodology through the vast speed-up
of the scouring process.

What did we learn?

By further tailoring our data sets,
getting more comprehensive data
and optimizing our model hyperparameters,
we can improve the performance and
utility of this method.

References

[1] Mlatticeabc: Generic Lattice Constant Prediction of Crystal Materials Using Machine Learning
Yuxin Li, Wenhui Yang, Rongzhi Dong, and Jianjun Hu
ACS Omega 2021 6 (17), 11585-11594
DOI: 10.1021/acsomega.1c00781

[2] Prediction of lattice constant in cubic perovskites
L.Q. Jiang, J.K. Guo, H.B. Liu, M. Zhu, X. Zhou, P. Wu, C.H. Li,
Journal of Physics and Chemistry of Solids, Volume 67, Issue 7, 2006, Pages 1531-1536,
ISSN 0022-3697,
https://doi.org/10.1016/j.jpcs.2006.02.004.

Thank you.