

# Materials Modeling Homework Assignment 5

Logan Kuhn

November 25, 2025

## Part I

### Water

Shown in figure 1 is the standard density with respect to time step. Upon taking the final density value, we can observe a final calculated density of 0.99713719. Likewise, shown in figure 2 is the density of water with respect to time step at 373 Kelvin. In this case, the final density value comes out to be 0.93631865, much different from the original calculated value at standard conditions. This reflects the boiling point of water well.

### Copper

Observing figure 3 will show the discontinuity at 1358 K, the spot we expect to see it at. This was generated without altering any conditions, but we are still able to see the discontinuity so long as we reduce the width of the plot line to a small enough value. Moreso, in figure 4 we can see timestamps of copper relinquishing its crystal structure as temperature increases.

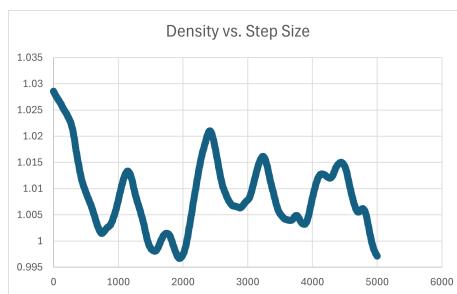


Figure 1: Water density under standard conditions

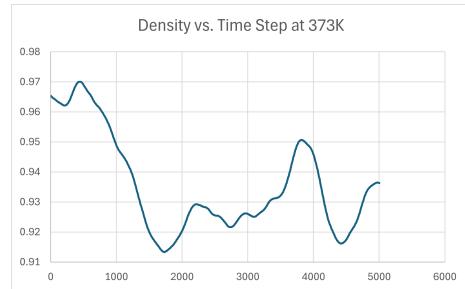


Figure 2: Water density at 373K

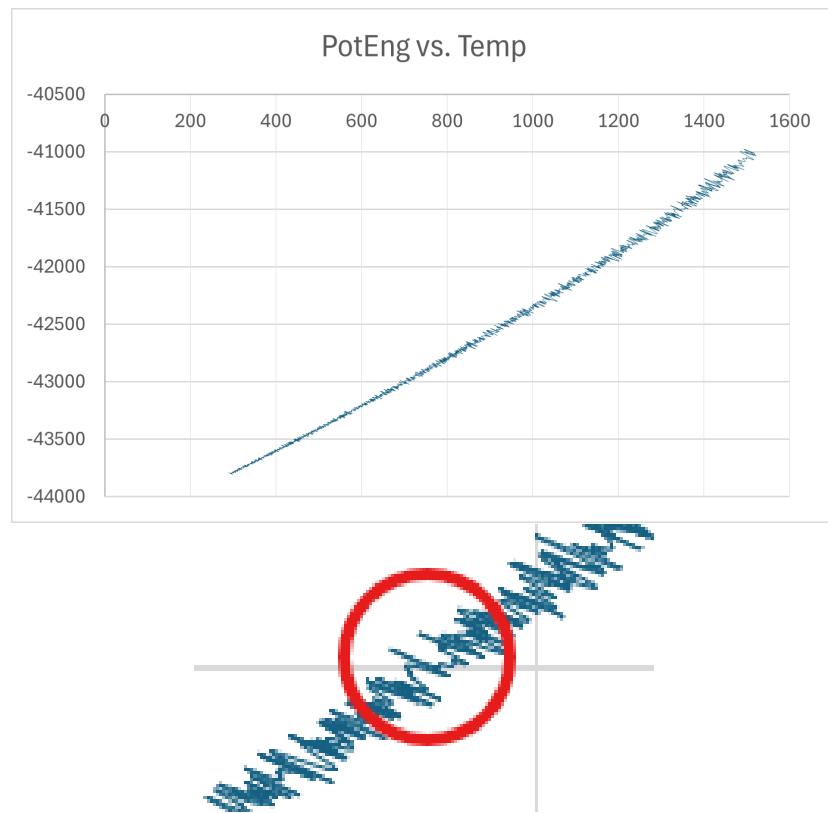


Figure 3: Discontinuity visualized in Copper at 1345

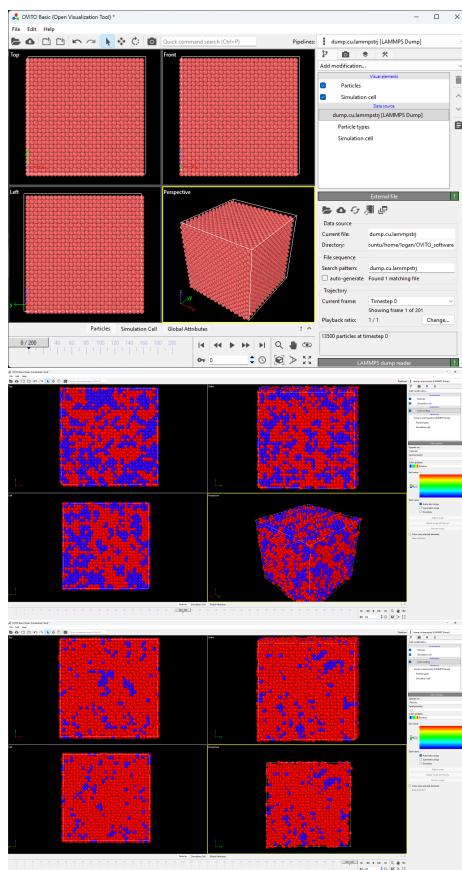


Figure 4: Structure of Copper at initial, half, and final time.

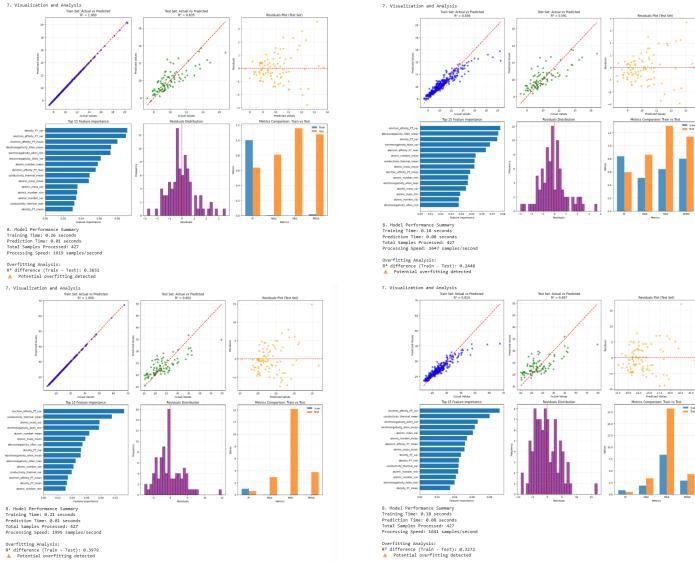


Figure 5: Comparison of DEED (top) and EFA (bottom) as well as GBR (left) and RFR (right)

## Part II

### Google Collab and Analysis

Comparing metrics between DEED and EFA as well as Radom Forrest Regression (RFR) and Gradient Boosting Regression (GBR), we can see the following performances in figure 5. We can see differences in, for example, the GBR shows a greater  $R^2$  value for Actual vs. Predicted training sets, where RFR shows much lower correlations. Another primary difference we see between DEED and EFA is in the residuals distribution and metrics comparison for training versus testing with  $R^2$ 's being higher for DEED than for EFA.

### DFT vs ML

Here we can analyze the differences between the Machine Learning models and that of DFT. We can see that differences arise, likely because machine learning is reliant on data sets and DFT input parameters. Errors in Machine Learning calculations could be a result of insufficient training data; whereas, errors in DFT could be a result in improper or inadequate input parameters. Shown in figure 6 is the differences in predicted energy per atom for reference.

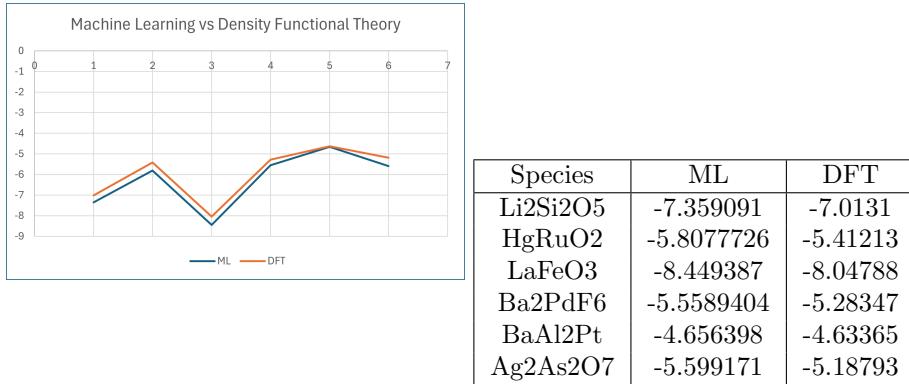


Figure 6: Machine Learning and DFT energy plotted against species

## Contributions

All work was completed independently with the exception of collaborating with classmates on rough concepts. Concepts deliberated on with classmates included how best to understand the interpretation of the Machine Learning algorithm and how to change target variables. All submitted work was created by myself.