Alonso Salcido, Chris Haub

Modeling Super Conductor Materials Using Lasso (L1 and Ridge (L2)

**Introduction**

Super conductors are materials that give little or no resistance to electrical current. The phenomenon is of interest in finding materials that allow electrical current to pass through with minimal loss. Benefits include reduced heat being generated as well as needing to use less electricity to power things in general. This is of interest in a wide variety of applications as the more conducive a material is the less money needs to be spent generating electricity to get to the source.

**Methods**

Two CSV files were provided for this case study. One was an entropy atomic mass CSV that had conduction rates while the other CSV was unique values for the materials. Two columns were dropped from the unique values CSV, critical temp and materials. Then the two data frames were merged into one file.

There were no missing values. The data was then scaled using the StandardScaler package from scikit-learn to subtract the mean value and produce a variance of 1 by dividing the standard deviation. 5 fold cross-validation was completed for 20 different strengths of regularization from -8 to 1. The best regularization strength was picked from the average out of fold mean squared error. The splits were randomly seeded.

Lastly, a final regularization strength was chose a 5 fold cross validation was repeated 100 times to test the out of fold mean squared error to produce an average mean squared error and select 95% Confidence Interval.

*Independent Variables*

number\_of\_elements mean\_atomic\_mass wtd\_mean\_atomic\_mass gmean\_atomic\_mass wtd\_gmean\_atomic\_mass entropy\_atomic\_mass wtd\_entropy\_atomic\_mass range\_atomic\_mass wtd\_range\_atomic\_mass std\_atomic\_mass wtd\_std\_atomic\_mass mean\_fie wtd\_mean\_fie gmean\_fie wtd\_gmean\_fie entropy\_fie wtd\_entropy\_fie range\_fie wtd\_range\_fie std\_fie wtd\_std\_fie mean\_atomic\_radius wtd\_mean\_atomic\_radius gmean\_atomic\_radius wtd\_gmean\_atomic\_radius entropy\_atomic\_radius wtd\_entropy\_atomic\_radius range\_atomic\_radius wtd\_range\_atomic\_radius std\_atomic\_radius wtd\_std\_atomic\_radius mean\_Density wtd\_mean\_Density gmean\_Density wtd\_gmean\_Density entropy\_Density wtd\_entropy\_Density range\_Density wtd\_range\_Density std\_Density wtd\_std\_Density mean\_ElectronAffinity wtd\_mean\_ElectronAffinity gmean\_ElectronAffinity wtd\_gmean\_ElectronAffinity entropy\_ElectronAffinity wtd\_entropy\_ElectronAffinity range\_ElectronAffinity wtd\_range\_ElectronAffinity std\_ElectronAffinity wtd\_std\_ElectronAffinity mean\_FusionHeat wtd\_mean\_FusionHeat gmean\_FusionHeat wtd\_gmean\_FusionHeat entropy\_FusionHeat wtd\_entropy\_FusionHeat range\_FusionHeat wtd\_range\_FusionHeat std\_FusionHeat wtd\_std\_FusionHeat mean\_ThermalConductivity wtd\_mean\_ThermalConductivity gmean\_ThermalConductivity wtd\_gmean\_ThermalConductivity entropy\_ThermalConductivity wtd\_entropy\_ThermalConductivity range\_ThermalConductivity wtd\_range\_ThermalConductivity std\_ThermalConductivity wtd\_std\_ThermalConductivity mean\_Valence wtd\_mean\_Valence gmean\_Valence wtd\_gmean\_Valence entropy\_Valence wtd\_entropy\_Valence range\_Valence wtd\_range\_Valence std\_Valence wtd\_std\_Valence

*Dependent Variable*

critical\_temp

**Results**

We find that regularization only begins to affect the mean squared error at large values of λ. Figure 1 as shown below, the mean squared error remains unchanged until λ exceeds 0.1.

Chart, bar chart

Description automatically generated

*Figure 1: Mean out of fold MSE across 100 runs while varying the strength of regularization. Mean Squared Error loss remains low until 0.1 or larger*

For the final λ=10­-8, the 100 runs produce an average Mean Squared Error of 0.3693

Chart, histogram

Description automatically generated

*Figure 2: The Mean Squared Error for out of fold predictions with a fixed strength of 0.00000001 (the best strength found in hyperparameter search).*

**Conclusion**

Our model produced mean squared error 0.3693. The best value of λ was found to be 10-8. Regularization strength produced very little effect below λ 0.1

**Appendix A**

Text

Description automatically generated

Text

Description automatically generated