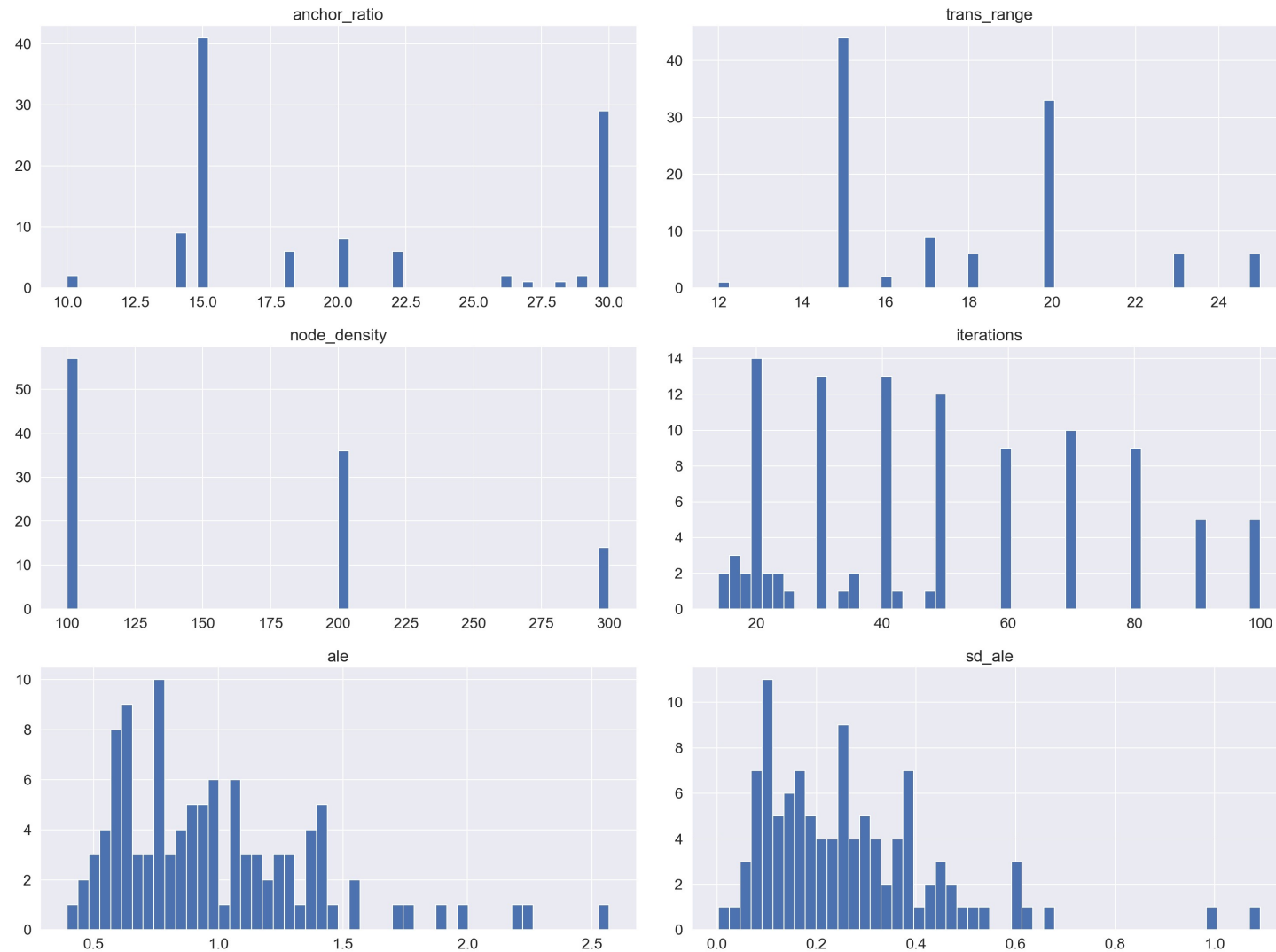


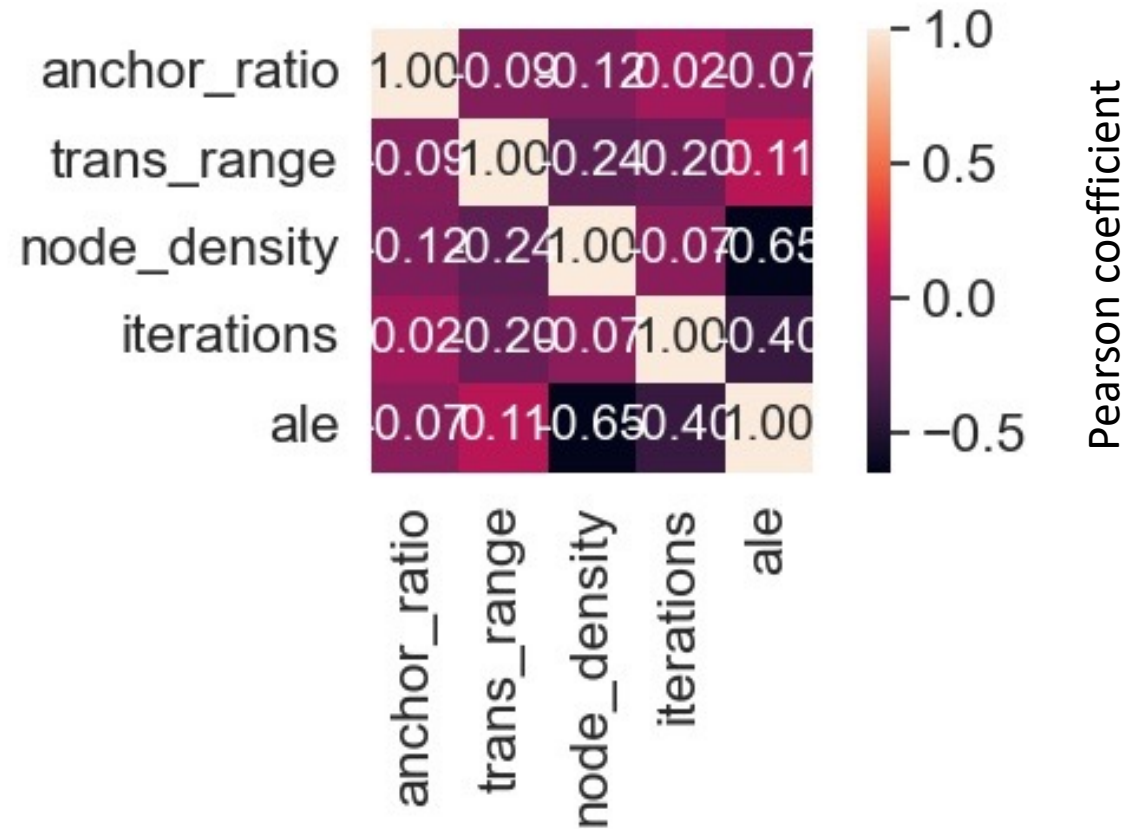
Part 2

Plots and results explanation

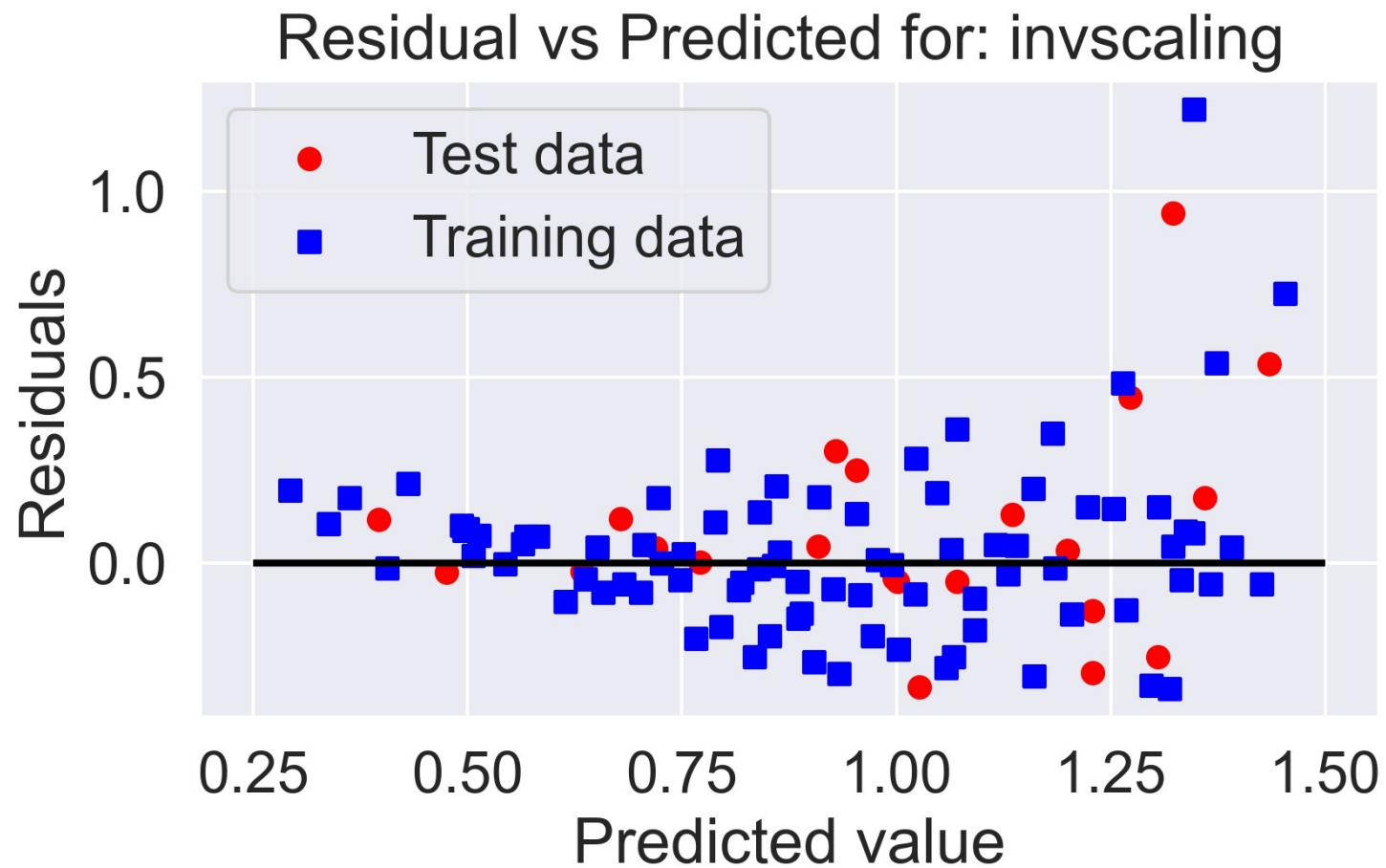
Histogram of feature variables



Pearson correlation



Residual vs Predicted plot



Explanation of results

1. In the case of perfect prediction, the residuals would be exactly 0. This will never be encountered in the real application scenario. However, in case of a good regression model the errors should be randomly distributed and there should not be a pattern. Detection of pattern in the residual plot would mean that model did not perform well in prediction.
2. For our case we see that, residuals are distributed along the center line ($y=0$). Hence the SGDRegressor performance was good. From the log file, you could see that MSE and R^2 for
 - Training data (MSE = 0.053, $R^2 = 0.654$)
 - Test data (MSE = 0.085, $R^2 = 0.568$)