Chart, line chart

Description automatically generated Chart, line chart

Description automatically generated

Lasso model sets equal or lower absolute value from Ridge model for each parameter, we can conclude that Ridge model causes weight decay with l2 regularization because it doesn’t “dispose” within calculations the unimportant features but still gives it weight higher than 0. As for the Lasso model, we see that it causes variable selection with l1 regularization (which include sparse solutions) as it ignores unimportant feature and gives less weight from Ridge model to some features.

Another behavior which indicates the differences between the models is that the Lasso features coefficient reaches 0 faster than Ridge model