untitled4-2

November 21, 2023

What is the optimal value of alpha for ridge and lasso regression? What will be the changes in the model if you choose double the value of alpha for both ridge and lasso? What will be the most important predictor variables after the change is implemented?

Answer : Best alpha value for Lasso : {'alpha': 0.001} - So optimal value - 100 Best alpha value for Ridge : {'alpha': 0.1} - so optimal value - 10

Increasing alpha in Ridge regression strengthens the regularization effect. It'll lead to more penalization of coefficients, likely shrinking them further towards zero. so R2 square value will drop.

Increasing alpha in Lasso intensifies regularization. Lasso tends to induce sparsity, driving more coefficients to absolute zero. So, doubling alpha means eliminating feautures. So R2 score will drop by nearly 0.9 %

Upon evaluating various models, it's evident that certain features play a pivotal role in elucidating the selling price of houses:

MSSubClass_90 (Type of Dwelling - DUPLEX - ALL STYLES AND AGES)

MSSubClass_120 (1-STORY PUD - Planned Unit Development - 1946 & NEWER)

RoofMatl_Membran (Roof Material - Membrane)

MSZoning_RL (Residential with Low Density - General Zoning Classification)

MSZoning_FV (Floating Village Residential)

MSZoning_RH (Residential High Density)

MSZoning_RM (Residential Medium Density)

Condition2_PosA (Adjacent to Positive Off-Site Feature)

RoofMatl_WdShngl (Roof Material - Wood Shingles)

Neighborhood_NoRidg (North Ridge)

OverallQual (Excellent Overall Quality) These features, ranging from dwelling types and zoning classifications to roofing materials and neighborhood designations, collectively contribute significantly to the explanation of house sale prices. Among them, 'OverallQual' specifically encapsulates the exceptional quality aspects of a property, encompassing superior craftsmanship, high-end materials, and an impressive overall aesthetic, which notably influences the property's market valuation.

You have determined the optimal value of lambda for ridge and lasso regression during the assignment. Now, which one will you choose to apply and why?

Answer : Opting for Lasso regression proves advantageous due to its intrinsic feature selection capability.It helps get rid of unnecessary stuff without losing what's important. This makes the model simpler and easier to understand while still staying accurate, like having a tidy and efficient space that works well for you.

After building the model, you realised that the five most important predictor variables in thelasso model are not available in the incoming data. You will now have to create another modelexcluding the five most important predictor variables. Which are the five most importantpredictor variables now?

Answer : MSSubClass_90, MSSubClass_120, TotalBsmtSF,OverallQual

How can you make sure that a model is robust and generalisable? What are the implications of the same for the accuracy of the model and why?

Answer :Ensuring a model is both robust and generalizable involves several practices:
a)Feature Selection and Regularization-> Employ methods like Lasso regularization that facilitate feature selection, emphasizing the most relevant predictors while disregarding less informative ones. This prevents overfitting and aids in creating a more adaptable model.

b)Avoid overfitting

c)We choose only the most important things for the model to look at and ignore the less helpful stuff. This helps the model understand better and not get confused by too many details.

Implications for Model Accuracy:

Strong vs. Perfect: Sometimes, a model that's really good at handling different situations might not be absolutely perfect with the data it was trained on. But that's okay because it means it's better at dealing with new things. Model accuracy should be $>$ 70-75% P-value of all the features is < 0.05 VIF of all the features are < 5

Working Everywhere: It's like making sure your solution works in many places rather than just being perfect in one specific spot. It might not be the best in that one place, but it's more useful overall.

Getting the Right Balance: We want our model to be good at its job but also flexible enough to handle new challenges. Finding this balance is important to make it really helpful in real life.

[ ]: