📉 Ridge Regression – Alpha Impact Analysis

This document analyzes how Ridge Regression performs on the California Housing dataset using different values of alpha (regularization strength).  
Ridge Regression applies L2 regularization, which shrinks coefficients toward zero but never eliminates them entirely, helping control model complexity.

# 🔧 What Ridge Does

- Ridge Regression adds a penalty to the cost function proportional to the square of the magnitude of coefficients.  
- It keeps all features but discourages large weights.  
- It's particularly useful when features are correlated or when overfitting needs to be controlled.

# 📊 Comparison: Ridge with alpha = 1 vs alpha = 10

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Alpha | R² Score | MAE | MSE | RMSE | Interpretation |
| 1.0 | 0.5758 | 0.5332 | 0.5559 | 0.7456 | Baseline Ridge (soft shrinkage, very similar to Linear) |
| 10 | 0.5761 | 0.5331 | 0.5555 | 0.7453 | Slightly improved generalization, smoother coefficients |

# 🧠 Coefficient Behavior

Ridge gently shrinks the coefficients as alpha increases. Even at alpha = 10, all features are retained.  
This is ideal when every feature carries some information, and we want to control the influence of each without eliminating them.

# 📌 Conclusion

Ridge Regression is a great tool for stabilizing a model when there's potential for overfitting or multicollinearity.  
In this dataset, alpha = 10 gave slightly better generalization than alpha = 1, making it a good candidate for regularized linear modeling.  
Unlike Lasso, Ridge keeps all features in the model while softening their impact.