PS8

Ahmed

04/04/2022

- 7) The dimension of your training data (housing_train) is 404 by 74 (N rows, N columns). There are 59 more X variables
- 8)

.metric	.estimator	.estimate	.config
rmse	standard	0.3898085	Preprocessor1_Model1
rsq	standard	NA	Preprocessor1_Model1

.metric	.estimator	mean	n	std_err	.config
rmse	standard	0.1435125	10	0.0098843	Preprocessor1_Model1

- 9)

.metric	.estimator	.estimate	.config
rmse	standard	0.2184031	Preprocessor1_Model1
rsq	standard	0.7163029	Preprocessor1_Model1

.metric	.estimator	mean	n	std_err	.config
rmse	standard	0.0933611	10	0.0072377	Preprocessor1_Model1

- 10) We cannot do such estimation with linear regression but we can using LASSO or Ridge. I don't think the results are good as we have high variance, RMSE of out sample is considerably different from the RMSE of in sample.