

Real Estate Price Prediction Using Elastic Net Regression Analysis

Bedrock Stat Analytics

February 17, 2019

Using Data file - NA.

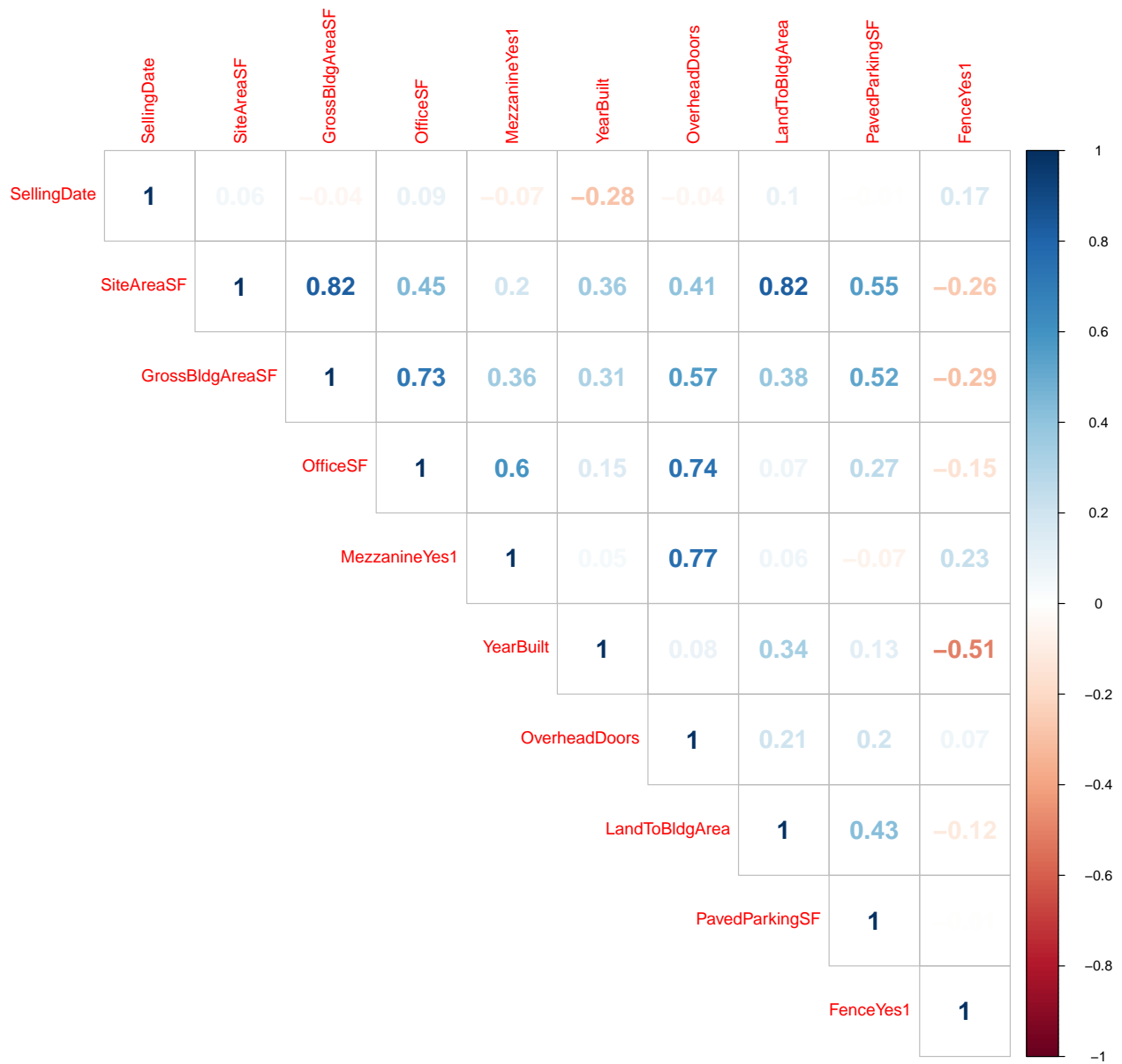
Summary Statistics

Description of df

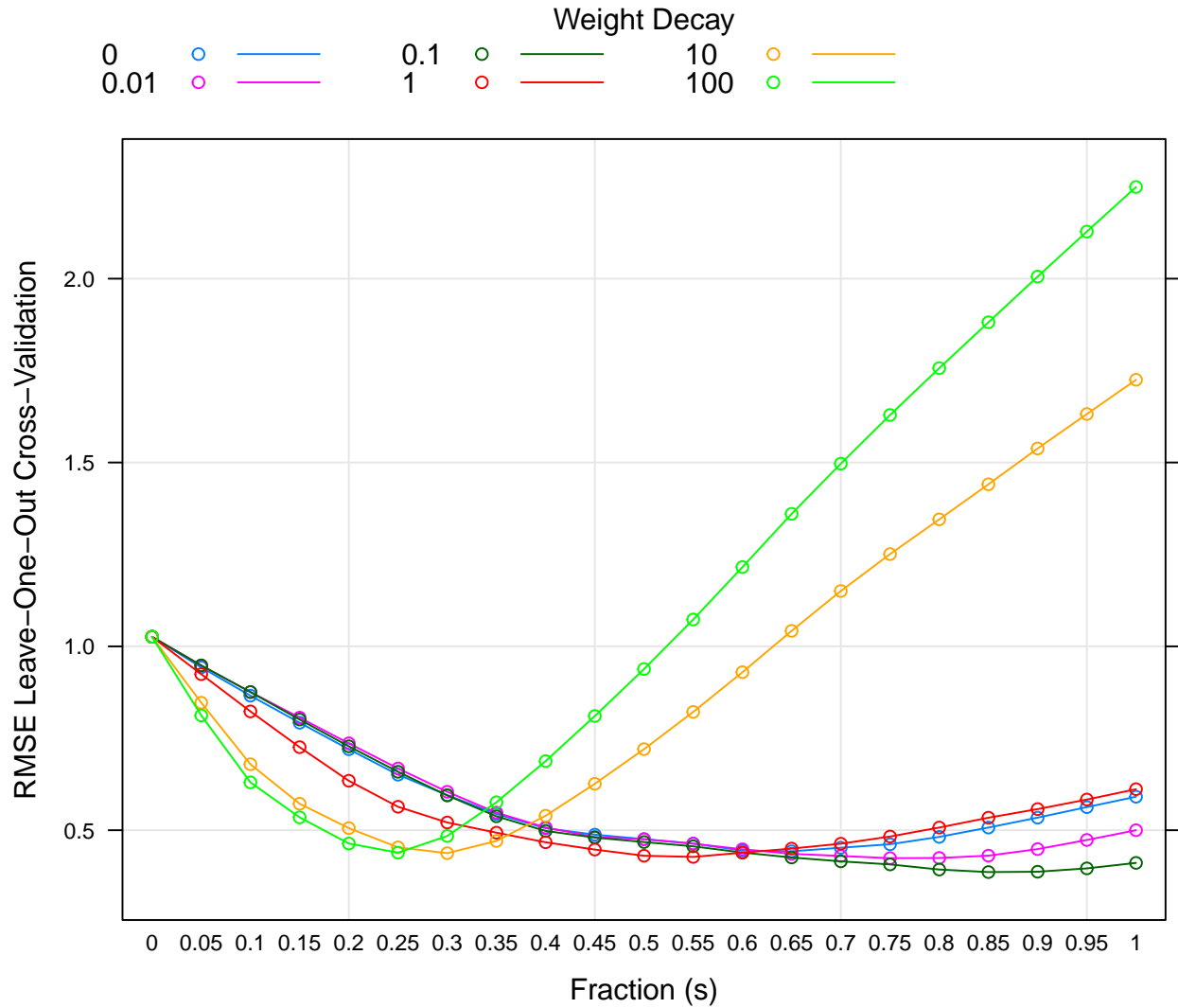
	N	Mean	Std. Dev.	Min	Max
SellingPrice	020	236307	125759	122750	450000
SellingDate	020	40473	1122	39006	42308
SiteAreaSF	020	18618	12944	6120	50094
GrossBldgAreaSF	020	4472	2045	2000	8400
OfficeSF	020	939	598	200	2400
MezzanineYes1	020	000	000	000	001
YearBuilt	020	1993	018	1948	2010
OverheadDoors	020	002	001	001	006
LandToBldgArea	020	003	001	001	007
PavedParkingSF	020	5453	4405	1800	17088
FenceYes1	020	000	000	000	001

Table 1: Summary statistics. Note that Mean and Std. Dev. are meaningless for categorical variables.

Correlations Between Predictors

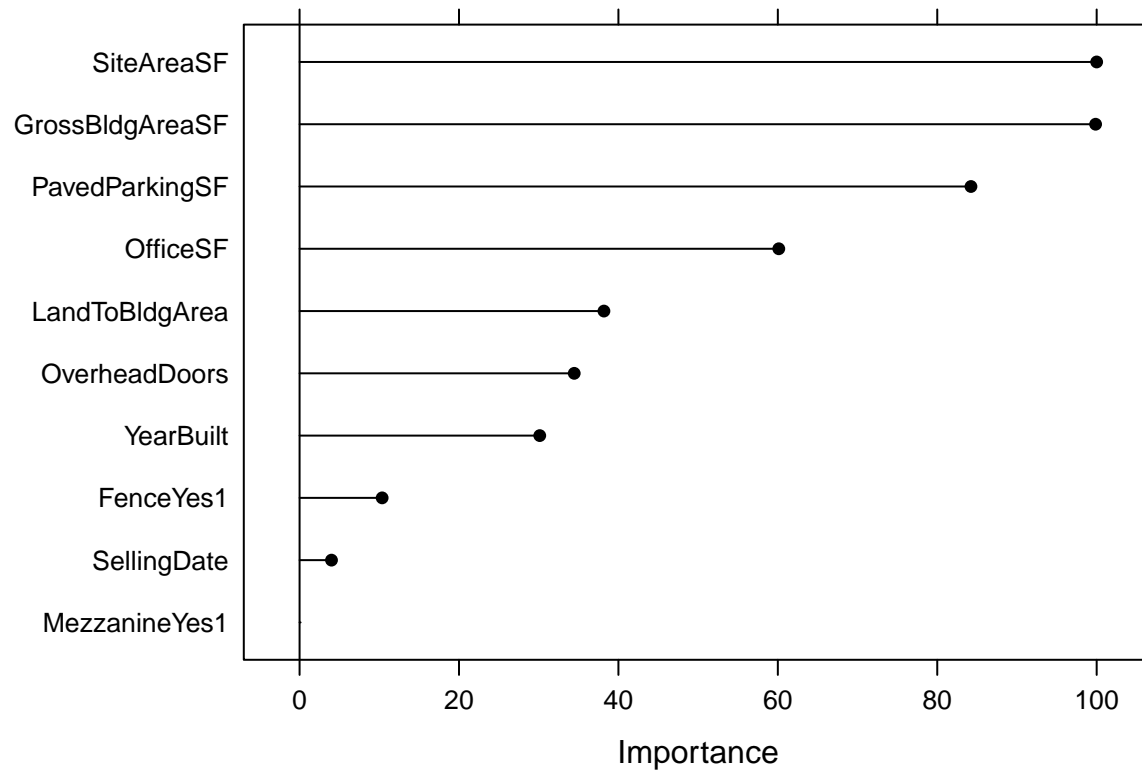


Tuning Parmenter Selection Using Leave-One-Out Cross Validation



From above plot, $\lambda = 0.10$ and $s = 0.85$ gives the minimum RMSE model.

Variable Importance



Rank of variables by Importance using Caret R Package. All measures of importance are scaled to have a maximum value of 100.

Standardized Model Coefficients

	Estimate
SellingDate	0.185
SiteAreaSF	0.362
GrossBldgAreaSF	0.338
OfficeSF	0.165
YearBuilt	0.177
OverheadDoors	0.000
LandToBldgArea	0.085
PavedParkingSF	0.110
MezzanineYes1	0.000
FenceYes1	-0.143

NOTE errors are not given for the predicted coefficients since it is yet unclear how meaningful the std. errors are in penalized regression.

Model Prediction

	Predicted Value	Prediction Error	R2
1	416816.820	0.386	0.845

Appendix

Table 2: Raw data table

SellingPrice	SellingDate	SiteAreaSF	GrossBldgAreaSF	OfficeSF	MezzanineYes1	YearBuilt	OverheadDoors	LandToBldgArea	PavedParkingSF	FenceYes1
128000	42123	8148	2976	1200	0	1948	1	2.737903	2500	0
149500	41760	11505	3200	1040	1	1980	3	3.595312	2500	1
163000	41808	10320	2880	547	0	1978	2	3.583333	4000	1
122750	40555	6120	3500	500	0	1961	2	1.748571	2400	0
255000	40787	14015	5400	1196	0	2010	1	2.595370	5450	0
169000	40669	16616	4000	480	0	1965	2	4.154000	12600	1
415000	40723	50094	7500	1000	0	1999	2	6.679200	6400	0
425000	42308	28314	6080	1080	0	2004	1	4.656908	1800	0
177500	40494	11205	3320	1320	0	2004	2	3.375000	5105	0
160000	39609	8160	2000	432	0	2006	1	4.080000	2500	0
425000	39623	43995	6000	800	0	2005	1	7.332500	13600	0
137500	39748	11505	3200	1040	1	1980	3	3.595312	2500	1
185000	39006	17028	5400	768	1	2000	2	3.153333	2500	0
139900	39123	21896	4000	200	0	2000	3	5.474000	4000	0
132000	39036	6950	2560	400	0	2006	1	2.714844	2420	0
152000	41893	6950	2560	400	0	2006	1	2.714844	2420	0
395000	41492	27442	6080	1080	0	2004	2	4.513487	17088	0
400000	39106	31973	8400	2400	1	2004	6	3.806309	8392	0
450000	40246	31973	8400	2400	1	2004	6	3.806309	8392	0
145000	39370	8160	2000	500	0	2007	1	4.080000	2500	0
NA	42852	37500	5000	1400	1	2004	5	7.500000	5800	0

Appendix