First Step. Upload Training Set CSV

Drag and drop file here
Limit 200MB per file • CSV
Browse files

training.csv
206.3KB

Select the target variable:

target

Select the variables to exclude:

image

Secod Step. Upload New Data for Predictions

Drag and drop file here
Limit 200MB per file • CSV
Browse files

test.csv
29.4KB

Select the variables to exclude::

image

Select column with path to images:

image

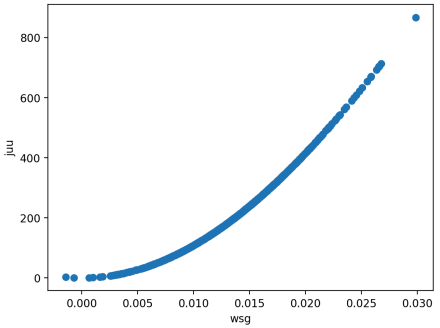
ML model Data Analysis Dashboard

Preprocessing features

Correlation between Features

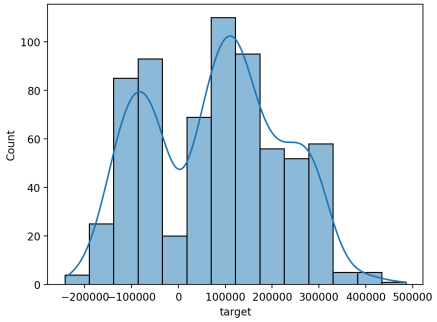
	Variable1	Variable2	Correlation	P-val
1	wsg	juu	0.9741	
7	target	date_day	0.1085	0.00
0	lux	fyt	0.1025	0.00
2	yyz	gox	0.0879	0.02
5	boz	lgh	0.0863	0.02
4	gox	foo	-0.0787	0.04
3	drt	gox	-0.0799	0.03
6	target	date_year	-0.0895	0.01

Correlated Features



Feature: wsg is redundant and will be removed

Target Variable



Column date is processed to get date_year, date_month, date_day and date_dayofweek

One Hot Encoding and Standard Transformation Training Set

	bar	baz	xgt	qgg	lux	yyz	drt	gox	foo	boz	fyt	lgh	hrt	juu	date_year	da
count	542	542	542	542	542	542	542	542	542	542	542	542	542	542	542	
mean	)0000000	)0000000	)0000000	)0000000	)0000000	)0000000	)00000000	)0000000	)0000000	)0000000	)0000000	)0000000	)0000000	)0000000	)0000000002	000
std	1.0009	1.0009	1.0009	1.0009	1.0009	1.0009	1.0009	1.0009	1.0009	1.0009	1.0009	1.0009	1.0009	1.0009	1.0009	
min	-0.8485	-0.9781	-3.6724	-2.5446	-2.5466	-2.5296	-21.2424	-2.638	-2.6292	-0.4022	-0.9602	-0.7317	-2.8845	-1.5443	-1.3198	
25%	-0.585	-0.9781	-0.6752	-0.5016	-0.7746	-0.5206	0.0924	-0.5515	-0.6862	-0.3565	-0.9602	-0.7317	-0.6993	-0.7941	-1.3198	
50%	-0.5302	-0.9781	0.009	-0.0058	-0.0613	0.0391	0.1056	0.0268	-0.0545	-0.2679	-0.9602	-0.7317	-0.0171	-0.1217	-0.188	
75%	0.228	1.0224	0.6769	0.4852	0.7775	0.5178	0.1071	0.5105	0.7095	-0.0541	1.0414	1.3666	0.6478	0.6003	0.9439	
max	3.1493	1.0224	3.2243	2.3757	1.7626	2.5662	0.1074	2.667	2.9665	10.6037	1.0414	1.3666	3.0341	4.6212	2.0758	

Running Model with all features

The model RandomForestRegressor(n_estimators=200, random_state=42) will be stored into the local directory.

Please write a name for your model file:

my_model_all_features

Save First Model

Model Evaluation

Model

Params: <bound method BaseEstimator.get_params of RandomForestRegressor(n_estimators=200, random_state=42)>

CV Training Results

CV RMSE: 41192.78

CV R2: 0.91

Rows in Set: 542

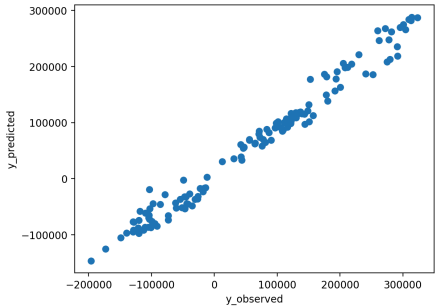
Testing Results

Test RMSE: 28666.52

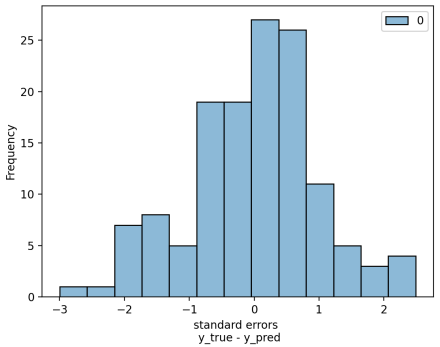
Test R2: 0.95

Rows in Set: 136

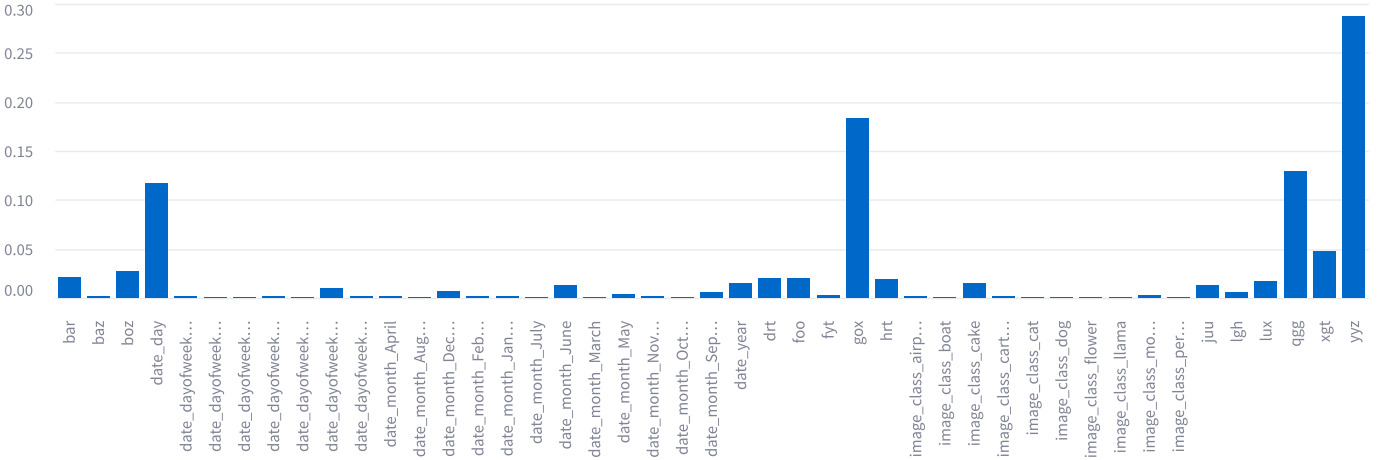
Test Obs vs Pred



Test Standard Errors



Feature Importances



Remember to click save model: name 'filename_v1' is not defined

Running Model with Selected Features

The model RandomForestRegressor(n_estimators=200, random_state=42) will be stored into the local directory.

Please write a name for your model file:

my_model_selected_features

Save Second Model

Model Evaluation

Selected Features for new model :['qgg', 'yyz', 'gox', 'date_day']

Model

Params: <bound method BaseEstimator.get_params of RandomForestRegressor(n_estimators=200, random_state=42)>

CV Training Results

CV RMSE: 53610.32

CV R2: 0.85

Rows in Set: 542

Testing Results

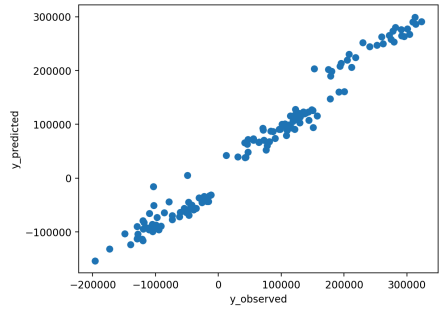
Test RMSE: 22890.36

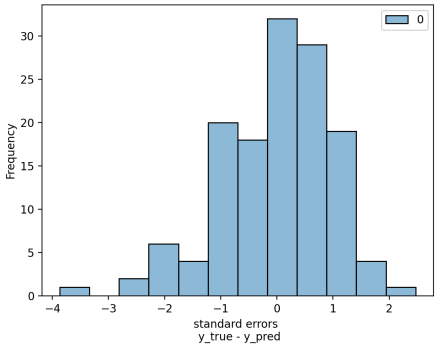
Test R2: 0.97

Rows in Set: 136

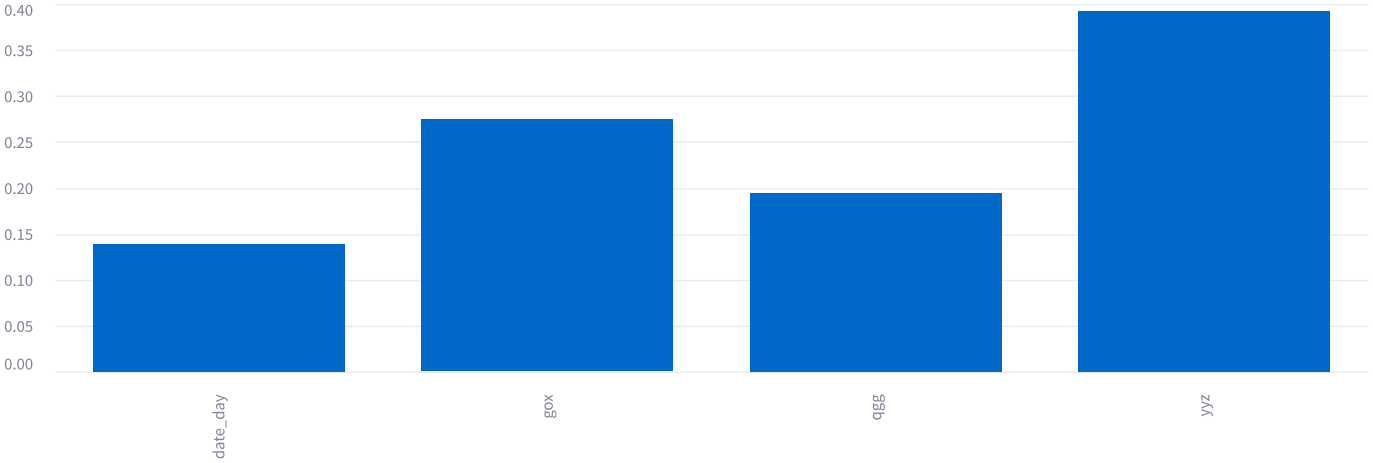
Test Obs vs Pred

Test Standard Errors





Feature Importances



Remember to click save model: name 'filename_v2' is not defined

Analize New Data Values

Upload your trained model in Pickle Format `.pkl` from your local computer.

Choose a `.pkl` file

Drag and drop file here

Limit 200MB per file • PKL

Browse files

my_model_selected_features_20250520_161327.pkl 2.5MB

File **'my_model_selected_features_20250520_161327.pkl'** uploaded successfully!

Model Details:

- Model Type: `RandomForestRegressor`
- Number of estimators: `200`
- Random state: `42`

A model is currently loaded in your session.

Clear Model

Predictions with New DataSet

The model doesn't predict image class, do you want to see the images from the uploaded newdata :

Select your answer:

- ○ Yes
- ◉ No

The predictions will be estimated.

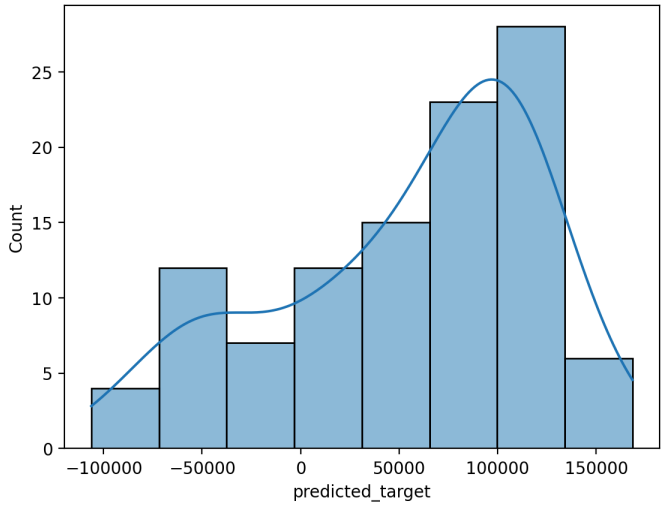
New Data Set One Hot Encoding and Standard Transformation

	bar	baz	xgt	qgg	lux	wsg	yyz	drt	gox	foo	boz	fyt	lgh	hrt	juu	date_y
count	107	107	107	107	107	107	107	107	107	107	107	107	107	107	107	
mean	)0000000	)0000000	)0000000	)0000000	)0000000	)0000000	)0000000	)0000000	)0000000	)0000000	)0000000	)0000000	)0000000	)0000000	)0000000	0.0000
std	1.0047	1.0047	1.0047	1.0047	1.0047	1.0047	1.0047	1.0047	1.0047	1.0047	1.0047	1.0047	1.0047	1.0047	1.0047	
min	-0.814	-1.0677	-2.3363	-2.472	-2.0292	-2.0927	-2.6763	-6.6896	-2.4209	-2.6964	-0.2722	-0.9907	-0.6532	-2.6933	-1.3688	
25%	-0.5847	-1.0677	-0.756	-0.4961	-0.8177	-0.5837	-0.4652	0.219	-0.3633	-0.6132	-0.2476	-0.9907	-0.6532	-0.5992	-0.6455	
50%	-0.5341	0.9366	-0.0487	-0.0619	-0.1751	-0.0439	-0.0121	0.2642	0.0691	-0.0182	-0.2194	-0.9907	-0.6532	0.0091	-0.1966	
75%	0.2206	0.9366	0.6866	0.5237	0.7169	0.5535	0.5247	0.2774	0.5322	0.8355	-0.1206	1.0094	1.5309	0.6576	0.4089	
max	3.1868	0.9366	2.4116	2.322	2.1098	2.8979	2.4599	0.2794	2.2394	2.3261	7.1318	1.0094	1.5309	2.5729	3.7215	

Predicted target

	predicted_target
0	116863.1979
1	18728.5405
2	-58821.4017
3	-36032.9047
4	-19499.9199

Predicted target Distribution



New Data Set with predictions

	image	date	bar	baz	xgt	qgg	lux	wsg	yyz	drt	gox	foo
0	test-img/boat/boat_3.jpg	2020-11-18	495.361	0	331.9667	21.0215	5190.7102	0.0152	-12.3441	-4.5872	-11.227	-7.9964
1	test-img/person/person_53.jpg	2019-01-20	127.4159	1	396.8519	5.6302	4767.2682	0.0108	5.4995	-25.6759	3.2398	23.3265
2	test-img/boat/boat_62.jpg	2019-07-05	-48.353	0	414.0968	-0.4063	4452.1813	0.0106	-0.7461	-15.2135	-0.5181	38.5688
3	test-img/car/car_7.jpg	2020-12-15	53.1638	0	365.5375	15.6114	5864.5716	0.0131	-17.03	-38.9334	14.9058	43.3215
4	test-img/cat/cat_2.jpg	2020-03-20	-32.5913	0	263.0537	-5.9384	4983.3841	0.0158	-7.0179	-7.0664	-25.4288	37.8503

Save predictions

Enter file name using your last_name-answer.csv

hernandez-answer.csv

Download CSV