# Sarit\_CrudeOilPrice

CS

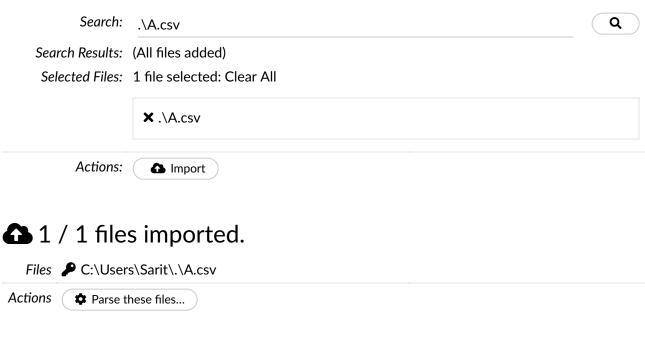
### Assistance

	Routine	Description
<b>C</b>	importFiles	Import file(s) into H <sub>2</sub> O
⊞	importSqlTable	Import SQL table into H <sub>2</sub> O
⊞	getFrames	Get a list of frames in H <sub>2</sub> O
×	splitFrame	Split a frame into two or more frames
Ø	mergeFrames	Merge two frames into one
	getModels	Get a list of models in H <sub>2</sub> O
	getGrids	Get a list of grid search results in $H_2O$
•	getPredictions	Get a list of predictions in H <sub>2</sub> O
₹≡	getJobs	Get a list of jobs running in H <sub>2</sub> O
4	runAutoML	Automatically train and tune many models
•	buildModel	Build a model
•	importModel	Import a saved model
•	predict	Make a prediction

# [ Import Files

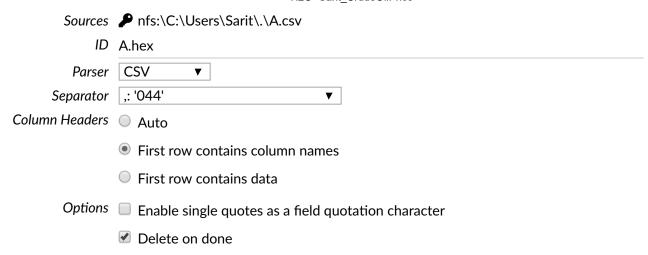
Setup Parse

127.0.0.1:54321/flow/index.html#



PARSE CONFIGURATION

1/17



#### EDIT COLUMN NAMES AND TYPES

Search	by	CO	lumn	name
--------	----	----	------	------

Seai	Cit by Columnia	116			
1	CrudeOilPri	Numeric ▼	25.55	24.91	24.78
2	ma_20	Numeric ▼	0.0	0.0	0.0
3	ma_100	Numeric ▼	0.0	0.0	0.0
4	std_20	Numeric ▼	0.0	0.0	0.0
5	std_100	Numeric ▼	0.0	0.0	0.0
6	rsi	Numeric ▼	0.0	0.0	0.0
7	bb_high_ind	Numeric ▼	0.0	0.0	0.0
8	bb_low_indi	Numeric ▼	0.0	0.0	0.0
9	26_ema	Numeric ▼	25.55	25.217692307692307	25.0604337111
10	12_ema	Numeric ▼	25.55	25.20333333333333	25.0381062355
11	macd	Numeric ▼	0.0	-0.01435897435897715	-0.0223274756
12	CCI	Numeric ▼	156.614458694477	156.614458694477	156.614458694
13	Index	Numeric ▼	1411.75	1413.5	1404.0
14	GoldPrice	Numeric ▼	282.7	281.1	281.4
15	SilverPrice	Numeric ▼	5.335	5.17	5.127
← Previous page → Next page					

Parse

### **≆** Job

Run Time 00:00:00.444

Remaining Time 00:00:00.0

Type Frame

Key Q A.hex

127.0.0.1:54321/flow/index.html# 2/17

Description Parse
Status DONE
Progress 100%
Done.
Actions Q View

### **⊞** A.hex

Actions: Wiew D		Build Model	Delete
Rows	Columns	Compressed Size	
5019	15	426KB	

#### **▼** COLUMN SUMMARIES

label	type	Missing	Zeros	+Inf	- Inf	min	max	mean	sigma
CrudeOilPrice	real	0	0	0	0	17.4500	145.2900	61.9786	26.3827
ma_20	real	0	19	0	0	0	139.5300	61.8246	26.4975
ma_100	real	0	99	0	0	0	126.0110	61.1384	26.8801
std_20	real	0	19	0	0	0	13.8067	2.1216	1.3258
std_100	real	0	99	0	0	0	27.8069	4.7927	3.7863
rsi	real	0	14	Θ	0	0	80.7686	51.0950	12.3874
bb_high_indicator	int	0	4775	0	0	0	1.0	0.0486	0.2151
bb_low_indicator	int	0	4776	0	0	Θ	1.0	0.0484	0.2147
26_ema	real	0	0	0	0	19.3583	138.0545	61.9059	26.1997
12_ema	real	0	0	0	0	18.7529	141.1385	61.9464	26.3061
macd	real	0	1	0	0	-9.8296	4.7026	0.0405	1.6332
CCI	real	0	Θ	0	0	-195.9657	156.6145	0.3892	66.9569
Index	real	0	Θ	0	0	676.0	3024.5000	1557.8188	559.6673
GoldPrice	real	0	0	0	0	255.1000	1888.7000	948.2849	475.2170
SilverPrice	real	0	0	0	0	4.0260	48.5840	14.9349	8.5601
4									•

← Previous 20 Columns

→ Next 20 Columns

▶ CHUNK COMPRESSION SUMMARY

▶ FRAME DISTRIBUTION SUMMARY

# **X** Split Frame

Frame: A.hex ▼

127.0.0.1:54321/flow/index.html# 3/17

 Splits:
 Ratio
 Key

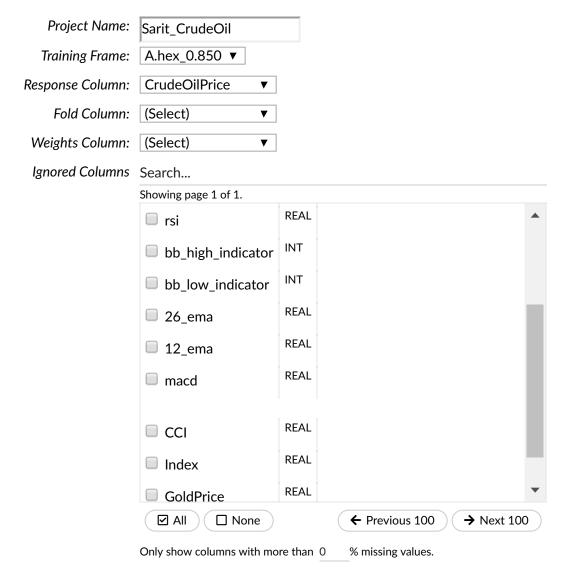
 0.85
 A.hex\_0.850
 X

 0.150
 A.hex\_0.150

 Add a new split
 Seed: 325035

### **⊞** Split Frames

### Run AutoML



127.0.0.1:54321/flow/index.html# 4/17

Validation Frame:	A.hex_0.150 ▼
Blending Frame:	(Select) ▼
Leaderboard Frame:	(Select) ▼
Balance classes:	
Exclude these algorithms	<b>☑</b> GLM
	✓ DRF
	□ GBM
	✓ XGBoost
	DeepLearning
	StackedEnsemble
Max models to build:	
Max Run Time (sec):	3600
Max Run Time Per Model (sec):	0
Early stopping metric:	AUTO ▼
Leaderboard sort metric:	AUTO ▼
Early stopping rounds:	3
Early stopping tolerance:	
nfolds:	5
Keep cross-validation predictions	<b>✓</b>
Keep cross-validation models	
Keep cross-validation fold assignment	
Seed:	-1
Checkpoints path:	

Build Model

# **≨**∃Job

Run Time 00:49:49.230

Remaining Time 00:00:00.0

Type Auto Model

Key Q Sarit\_CrudeOil

127.0.0.1:54321/flow/index.html# 5/17

Description AutoML build

Status DONE

Progress 100%

Done.

Actions Q View

### **₹** Leaderboard



#### ▼ MODELS

```
models sorted in order of mean_residual_deviance, best first
    model_id
                                                          mean_residual_deviance
                                                                                   rmse
                                                                                   0.52275
0
    StackedEnsemble_AllModels_AutoML_20190829_145323
                                                          0.273272964022179
    GBM_grid_1_AutoML_20190829_145323_model_19
1
                                                          0.3168315278608414
                                                                                   0.56287
2
    GBM_grid_1_AutoML_20190829_145323_model_34
                                                          0.3361789672132366
                                                                                   0.57980
3
    GBM_grid_1_AutoML_20190829_145323_model_21
                                                          0.34613364873822017
                                                                                   0.58833
    GBM_3_AutoML_20190829_145323
                                                          0.34779547542479533
                                                                                   0.58974
4
    GBM_2_AutoML_20190829_145323
                                                          0.35734706884524436
                                                                                   0.59778
5
    GBM_grid_1_AutoML_20190829_145323_model_3
                                                          0.36998642914187563
                                                                                   0.60826
6
    GBM_grid_1_AutoML_20190829_145323_model_42
7
                                                          0.3730351344230987
                                                                                   0.61076
    GBM_1_AutoML_20190829_145323
                                                          0.3754047359704281
8
                                                                                   0.61270
9
    GBM_4_AutoML_20190829_145323
                                                          0.3877439707891518
                                                                                   0.62269
   GBM_grid_1_AutoML_20190829_145323_model_46
                                                          0.39548511854033314
                                                                                   0.62887
10
11 GBM_grid_1_AutoML_20190829_145323_model_41
                                                          0.3984072982641279
                                                                                   0.63119
   GBM_grid_1_AutoML_20190829_145323_model_43
                                                          0.40738442390851665
                                                                                   0.63826
12
   GBM_grid_1_AutoML_20190829_145323_model_7
                                                          0.41524486707617314
                                                                                   0.64439
14 GBM_grid_1_AutoML_20190829_145323_model_2
                                                          0.4183937996166078
                                                                                   0.64683
    GBM_grid_1_AutoML_20190829_145323_model_20
15
                                                          0.42266935333489203
                                                                                   0.65013
   GBM_grid_1_AutoML_20190829_145323_model_47
                                                          0.44077190486900825
                                                                                   0.66390
17 GBM_grid_1_AutoML_20190829_145323_model_35
                                                          0.46071201818468877
                                                                                   0.67875
   GBM_grid_1_AutoML_20190829_145323_model_10
18
                                                          0.47380238029304655
                                                                                   0.68833
   GBM_grid_1_AutoML_20190829_145323_model_17
                                                          0.4968359913852461
                                                                                   0.70486
   GBM_grid_1_AutoML_20190829_145323_model_49
                                                          0.4984843369608376
                                                                                   0.70603
20
   GBM_grid_1_AutoML_20190829_145323_model_44
                                                          0.4989676088057719
                                                                                   0.70637
    GBM_grid_1_AutoML_20190829_145323_model_38
                                                          0.5027101832047299
                                                                                   0.70902
```

#### ▼ EVENT LOG

#### Actions taken and discoveries made by AutoML

	timestamp	level	stage	message	
0	14:53:23.552	Info	Workflow	Project: Sarit_CrudeOil	
1	14:53:23.553	Info	Workflow	AutoML job created: 2019.08.29 14:53:23.543	
2	14:53:23.555	Info	ModelTraining	Disabling Algo: XGBoost as requested by the user	

127.0.0.1:54321/flow/index.html# 6/17

				<del>-</del>	
	timestamp	level	stage	message	
3	14:53:23.556	Info	ModelTraining	Disabling Algo: DRF as requested by the user.	
4	14:53:23.556	Info	ModelTraining	Disabling Algo: GLM as requested by the user.	
5	14:53:23.556	Info	Workflow	Build control seed: -1 (random)	
6	14:53:23.558	Info	DataImport	training frame: Frame key: automl_training_A.hex_0.850 cols: 15 rows: 4269 chunks: 32 size: 376656 checksum: -2319176461513032855	
7	14:53:23.558	Info	DataImport	validation frame: Frame key: A.hex_0.150 cols: 1 rows: 750 chunks: 32 size: 95364 checksum: -7314559962142721135	
8	14:53:23.558	Info	DataImport	leaderboard frame: NULL	
9	14:53:23.559	Info	DataImport	response column: CrudeOilPrice	
10	14:53:23.559	Info	DataImport	fold column: null	
11	14:53:23.559	Info	DataImport	weights column: null	
12	14:53:23.675	Info	Workflow	Setting stopping tolerance adaptively based on the training frame: 0.015305126468886985	
13	14:53:23.693	Info	Workflow	AutoML build started: 2019.08.29 14:53:23.693	
14	14:53:23.722	Debug	ModelTraining	GBM 1 started	
15	14:53:33.275	Debug	ModelTraining	GBM 1 complete	•
4				<b>•</b>	

### **₩** Model

Model ID: StackedEnsemble\_BestOfFamily\_AutoML\_20190829\_145323 Algorithm: Stacked Ensemble Actions: **2** Refresh **7** Predict... **▲** Download POJO **▲** Download Model Deployment Package (MOJO) ■ Export **≡** Inspect ♣ Download Gen Model ▶ MODEL PARAMETERS ▶ OUTPUT

NAMES

▶ COLUMN\_TYPES

DOMAINS

▼ OUTPUT - TRAINING\_METRICS

7/17 127.0.0.1:54321/flow/index.html#

```
H2O- Sarit_CrudeOilPrice
                 model StackedEnsemble_BestOfFamily_AutoML_20190829_145323
        model_checksum 2003951472482220516
                 frame automl_training_A.hex_0.850
        frame_checksum -2319176461513032855
           description ·
        model_category Regression
          scoring_time 1567064536492
           predictions ·
                   MSE 0.209256
                  RMSE 0.457445
                  nobs 4269
    custom_metric_name .
   custom_metric_value 0
                    r2 0.999699
mean_residual_deviance 0.209256
                   mae 0.313442
                 rmsle 0.009407
     residual_deviance 893.313913
         null_deviance 2971503.323504
```

null\_degrees\_of\_freedom 4268 residual\_degrees\_of\_freedom 4266

AIC 5445.340883

```
▼ OUTPUT - VALIDATION_METRICS
                      model metalearner_AUTO_StackedEnsemble_BestOfFamily_AutoML_20190829_145
             model_checksum 5651813717943734712
                      frame
             frame_checksum 0
                description .
             model_category Regression
               scoring_time 1567064535262
                predictions ·
                        MSE 0.396255
                       RMSE 0.629488
                       nobs 750
         custom_metric_name .
        custom_metric_value 0
                         r2 0.999430
     mean_residual_deviance 0.396255
                        mae 0.418694
                      rmsle 0.011953
          residual_deviance 297.191284
              null deviance 521254.048585
                        AIC 1442.134884
    null_degrees_of_freedom 749
residual_degrees_of_freedom 747
```

127.0.0.1:54321/flow/index.html# 8/17

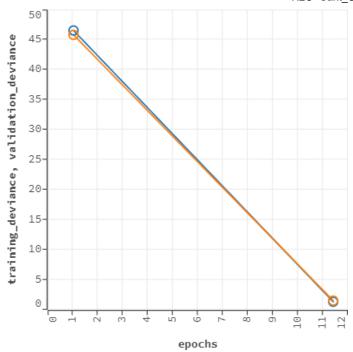
```
▼ OUTPUT - CROSS_VALIDATION_METRICS
```

```
model StackedEnsemble_BestOfFamily_AutoML_20190829_145323
             model checksum 2003951472482220516
                      frame levelone_training_StackedEnsemble_BestOfFamily_AutoML_20190829_14
             frame_checksum 1002510258213805427
                            5-fold cross-validation on training data (Metrics computed for
                description
                            combined holdout predictions)
             model_category Regression
               scoring_time 1567064535270
               predictions ·
                       MSE 0.503108
                       RMSE 0.709301
                       nobs 4269
         custom_metric_name .
        custom_metric_value 0
                        r2 0.999277
    mean_residual_deviance 0.503108
                       mae 0.483127
                      rmsle 0.015282
          residual_deviance 2147.770176
             null_deviance 2971655.768073
                       AIC 9190.310075
   null_degrees_of_freedom 4268
residual_degrees_of_freedom 4266
```

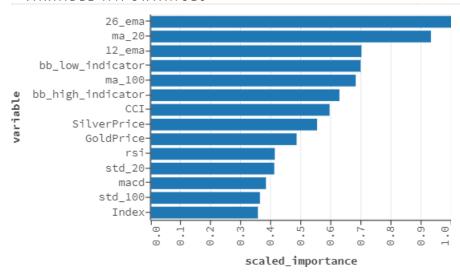
#### ▶ OUTPUT - HELP



127.0.0.1:54321/flow/index.html# 9/17



#### ▼ VARIABLE IMPORTANCES



- ▶ OUTPUT
- ▶ COLUMN\_TYPES
- ▶ CROSS\_VALIDATION\_MODELS
- ▶ CROSS\_VALIDATION\_PREDICTIONS
- ▶ OUTPUT STATUS OF NEURON LAYERS (PREDICTING CRUDEOILPRICE, REGRESSION, GAUSSIAN DISTRIBUTION, QUADRATIC LOSS, 381 WEIGHTS/BIASES, 11.2 KB, 48,705 TRAINING SAMPLES, MINI-BATCH SIZE 1)
- ▶ OUTPUT SCORING HISTORY
- ▶ OUTPUT TRAINING\_METRICS
- ▶ OUTPUT VALIDATION\_METRICS

- ▶ OUTPUT CROSS\_VALIDATION\_METRICS
- ▶ OUTPUT CROSS-VALIDATION METRICS SUMMARY
- ▶ OUTPUT VARIABLE IMPORTANCES
- ▼ PREVIEW POJO

Preview POJO

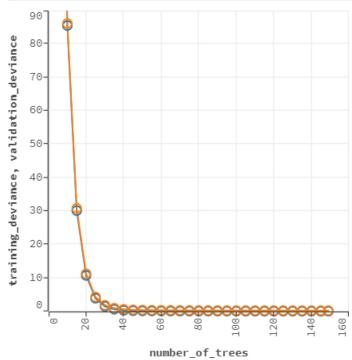


Model ID: GBM\_3\_AutoML\_20190829\_145323

Algorithm: Gradient Boosting Machine

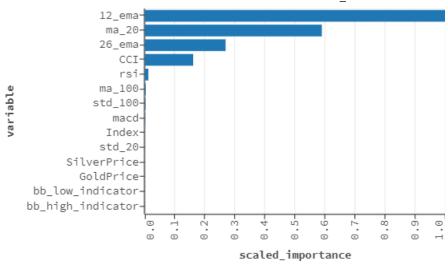
#### ▶ MODEL PARAMETERS

#### ▼ SCORING HISTORY - DEVIANCE



#### ▼ VARIABLE IMPORTANCES

127.0.0.1:54321/flow/index.html# 11/17



- ▶ OUTPUT
- ▶ COLUMN\_TYPES
- ▶ CROSS\_VALIDATION\_MODELS
- ▶ CROSS\_VALIDATION\_PREDICTIONS
- ▶ OUTPUT MODEL SUMMARY
- ▶ OUTPUT SCORING HISTORY
- ▶ OUTPUT TRAINING\_METRICS
- ▶ OUTPUT VALIDATION\_METRICS
- ▶ OUTPUT CROSS\_VALIDATION\_METRICS
- ▶ OUTPUT CROSS-VALIDATION METRICS SUMMARY
- ▶ OUTPUT VARIABLE IMPORTANCES
- ▼ PREVIEW POJO
- Preview POJO

### **Predict**

Name: prediction-2d89cf63-558

Model: StackedEnsemble\_BestOfFamily\_AutoML\_20190829\_145323

Frame: A.hex\_0.150

Actions: Predict

### **7** Prediction

127.0.0.1:54321/flow/index.html# 12/17

▼

Actions: ( **≡** Inspect

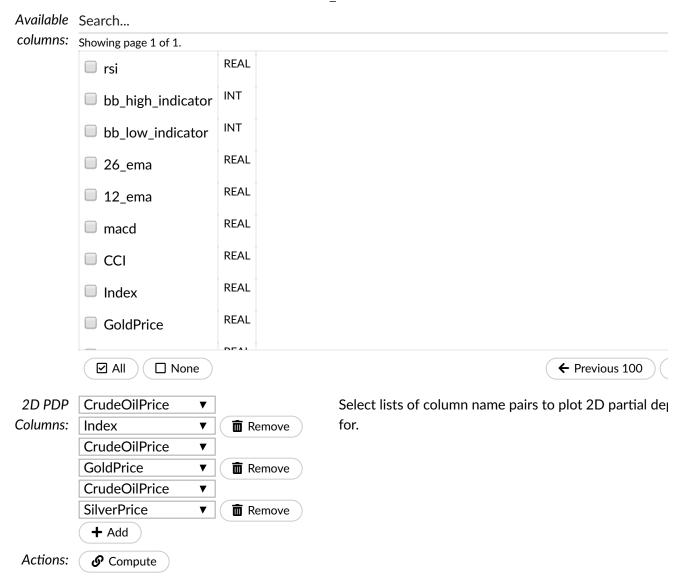
#### **▼ PREDICTION**

```
model StackedEnsemble_BestOfFamily_AutoML_20190829_145323
             model_checksum 2003951472482220516
                      frame A.hex_0.150
             frame_checksum -7314559962142721135
                description ·
             model_category Regression
               scoring_time 1567065170746
                predictions prediction-2d89cf63-558f-4053-9b91-2859bdf52654
                        MSE 0.396255
                       RMSE 0.629488
                       nobs 750
         custom\_metric\_name .
        custom_metric_value 0
                         r2 0.999430
    mean_residual_deviance 0.396255
                        mae 0.418694
                      rmsle 0.011953
          residual_deviance 297.191284
              null_deviance 521254.048585
                        AIC 1442.134884
   null_degrees_of_freedom 749
residual_degrees_of_freedom 747
E Combine predictions with frame
```

## Partial Dependence

Save Destination PDP as:	pdp-77f431d8-a32b-4a6f-a973-4b5539	
Model:	StackedEnsemble_BestOfFamily_AutoMI	_20190829_145323
Frame:	A.hex_0.850	
row_index	-1	Row for which partial dependence will be calculated in whole input frame (-1 for all).
nbins	20	How many levels should PDP compute. More levels w slower.
Select columns?		Checking this will allow you to select custom columns default, the top 10 features are used. Those features a variable importance.

127.0.0.1:54321/flow/index.html# 13/17



### **₹** Job

Run Time 00:08:01.167

Remaining Time 00:00:00.0

Type PartialDependence

Key Q pdp-77f431d8-a32b-4a6f-a973-4b5539790a02

Description PartialDependence

Status DONE

Progress 100%

Done.

Actions Q View

### **⊞** Partial Dependence Summary

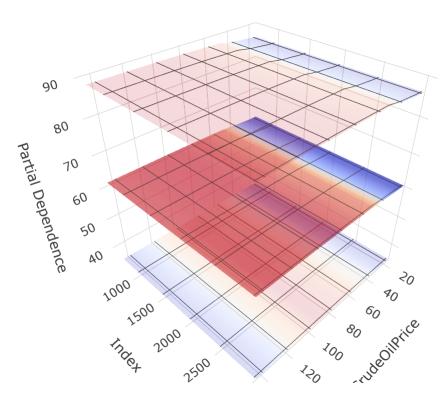
Model ID: StackedEnsemble\_BestOfFamily\_AutoML\_20190829\_145323

127.0.0.1:54321/flow/index.html#

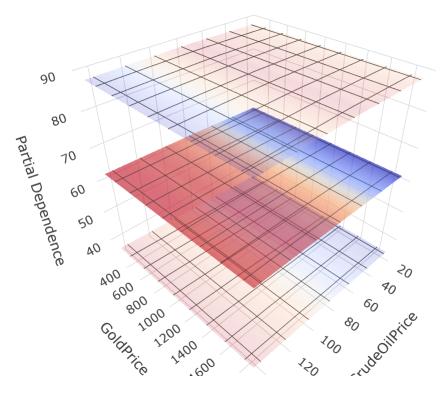
### Frame ID: A.hex\_0.850

Show PDP Data Table?

#### ▼ CRUDEOILPRICE VS INDEX

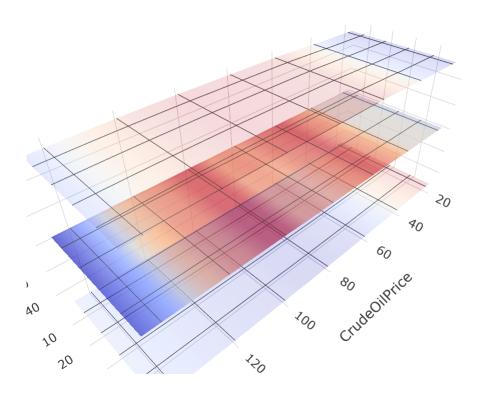


#### ▼ CRUDEOILPRICE VS GOLDPRICE



#### ▼ CRUDEOILPRICE VS SILVERPRICE

127.0.0.1:54321/flow/index.html# 15/17



# Partial Dependence

Save	pdp-5ab464cb-50f6-43ae-b83c-3b9889	
Destination		
PDP as:		
Model:	StackedEnsemble_BestOfFamily_AutoM	L_20190829_145323 ▼
Frame:	A.hex_0.150	
row_index	-1	Row for which partial dependence will be calculated in
		whole input frame (-1 for all).
nbins	20	How many levels should PDP compute. More levels w
		slower.
Select	•	Checking this will allow you to select custom columns
columns?		default, the top 10 features are used. Those features a
		variable importance.

127.0.0.1:54321/flow/index.html# 16/17

			1.20 04.10.4450 1.60				
Available	Search						
columns:	Showing page 1 of 1.						
	☐ CrudeOilPrice	REAL					
	□ ma_20	REAL					
	☐ ma_100	REAL					
	□ std_20	REAL					
	std_100	REAL					
	rsi rsi	REAL					
	☐ bb_high_indicator	INT					
	☐ bb_low_indicator	INT					
	☐ 26_ema	REAL					
		RΕΔΙ					
	☑ All ☐ None		← Previous 100				
2D PDP	+ Add		Select lists of column name pairs to plot 2D partial de				
Columns:			for.				
Actions:	Compute						

127.0.0.1:54321/flow/index.html#