PGP SL Eff Prod 2

Assistance

	Routine	Description
	`importFiles`	Import file(s) into H_2O
	`importSqlTable`	Import SQL table into H_2O
	`getFrames`	Get a list of frames in H_2O
	`splitFrame`	Split a frame into two or more frames
	`mergeFrames`	Merge two frames into one
	`getModels`	Get a list of models in H_2O
	`getGrids`	Get a list of grid search results in H_2O
	`getPredictions`	Get a list of predictions in H_2O
	`getJobs`	Get a list of jobs running in H_2O
	`runAutoML`	Automatically train and tune many models
	`buildModel`	Build a model
	`importModel`	Import a saved model
	`predict`	Make a prediction

Import Files

Search: C:\Users\ritwi\Desktop\ripik ai\Sri Lanka Efficiency Use Case\product 2\flatfile_prod2.csv

Search Results: (All files added)

Selected Files: 1 file selected: Clear All

✖ C:\Users\ritwi\Desktop\ripik ai\Sri Lanka Efficiency Use Case\product 2\flatfile_prod2.csv

Actions: Import

1 / 1 files imported.

Files C:\Users\ritwi\Desktop\ripik ai\Sri Lanka Efficiency Use Case\product 2\flatfile_prod2.csv

Actions Parse these files...

Setup Parse

PARSE CONFIGURATION

Sources nfs:\C:\Users\ritwi\Desktop\ripik ai\Sri Lanka Efficiency Use Case\product 2\flatfile_prod2.csv

ID flatfile_prod2.hex

Parser CSV

Separator ,: '044'

Escape Character 0

Column Headers
- ○ Auto
- ◉ First row contains column names
- ○ First row contains data

Options
- ☐ Enable single quotes as a field quotation character
- ☑ Delete on done

EDIT COLUMN NAMES AND TYPES

Search by column name...

1	Date	Time	20-Dec-21	21-Dec-21	22-Dec-21	23-Dec-21	24-Dec-21	25-Dec-21
2	P5_eff	Numeric	89.8	93	93.1	93.1	93	91.3
3	Adj_eff	Numeric	85.30259366	83.83561644	85.49089611	83.17460317	87.86798447	90.0615085
4	P5_sap_dra	Numeric	57.01	56.72	56.74	56.68	56.61	56.89
5	P5_actual_	Numeric	57	56.7	56.7	56.7	56.6	56.9
6	P5_seed	Numeric	9	6	9	12	17	16
7	Total SAP I	Numeric	313.15	306.91	310.52	310.23	308.79	309.52
8	Adjusted D	Numeric	312.3	306.6	280.1	302.4	283.3	308.9
9	Total Cull	Numeric	30	30	30	30	30	30
10	Melting co:	Numeric	10242	10193	10122	10256	10232	10259
11	Barrier Bo	Numeric	6324	5468	4640	6564	5467	5452
12	Capacitive	Numeric	31475	30024	30024	30884	29814	30484
13	Boosting (I	Numeric	37799	35492	34664	37448	35281	35936
14	Barrier Bo	Numeric	20.19	17.82	14.94	21.16	17.7	17.61
15	Capacitive	Numeric	100.51	97.83	96.69	99.55	96.55	98.49

Previous page | Next page

Parse

Job

Run Time 00:00:00.96
Remaining Time 00:00:00.0
Type Frame
Key flatfile_prod2.hex
Description Parse
Status DONE
Progress 100%
Done.
Actions View

flatfile_prod2.hex

Actions: View Data | Split | Build Model | Run AutoML | Predict | Download | Delete | Export

Rows	*Columns*	*Compressed Size*
33	165	35KB

▸ COLUMN SUMMARIES

▸ CHUNK COMPRESSION SUMMARY

▸ FRAME DISTRIBUTION SUMMARY

Split Frame

Frame: flatfile_prod2.hex

Splits:	Ratio	Key	
	0.7	flatfile_prod2.hex_0.70	✖
	0.30	flatfile_prod2.hex_0.30	
	Add a new split		
Seed:	54		

Create

Split Frames

Type	Key	Ratio
	flatfile_prod2.hex_0.70	0.7
	flatfile_prod2.hex_0.30	0.30000000000000004

Build a Model

Select an algorithm: Distributed Random Forest

PARAMETERS

model_id	drf-b5713458-ede7-46c5-b920-b5ecbea620c7	Destination id for this r auto-generated if not s
training_frame	flatfile_prod2.hex	Id of the training data f
validation_frame	flatfile_prod2.hex	Id of the validation data
nfolds	0	Number of folds for K-f cross-validation (0 to di >= 2).
response_column	P5_eff	Response variable colu
ignored_columns	Search...	Names of columns to ig training.

Showing page 1 of 17. -162 ignored.

☑ Date	TIME	3% NA
☑ P5_eff	REAL	3% NA
☑ Adj_eff	REAL	3% NA
☐ P5_sap_draw	REAL	3% NA
☐ P5_actual_draw	REAL	3% NA
☐ P5_seed	INT	3% NA
☐ Total SAP Draw	REAL	3% NA
☐ Adjusted Draw	REAL	3% NA
☐ Total Cullet %	INT	3% NA
☐ Melting cost-Rs/ton/Draw	INT	3% NA

☑ All ☐ None ← Previous 10 → Next 10

Only show columns with more than 0 % missing values.

ignore_const_cols	☑	Ignore constant column
ntrees	10	Number of trees.
max_depth	20	Maximum tree depth (0 unlimited).
min_rows	1	Fewest allowed (weight observations in a leaf.

nbins	20	For numerical columns build a histogram of (at this many bins, then spl best point
seed	-1	Seed for pseudo randor number generator (if ap
mtries	-1	Number of variables ra sampled as candidates split. If set to -1, defaul sqrt{p} for classification for regression (where p of predictors
sample_rate	0.632	Row sample rate per tre 0.0 to 1.0)

ADVANCED

score_each_iteration	☐	Whether to score durin iteration of model train
score_tree_interval	0	Score the model after e many trees. Disabled if
fold_column	(Choose...)	Column with cross-valic fold index assignment p observation.
offset_column	(Choose...)	Offset column. This will added to the combinati columns before applyin function.
weights_column	(Choose...)	Column with observatic weights. Giving some observation a weight of equivalent to excluding the dataset; giving an observation a relative w 2 is equivalent to repea row twice. Negative we not allowed. Note: Weig per-row observation we and do not increase the the data frame. This is t the number of times a r repeated, but non-integ values are supported as During training, rows w higher weights matter r due to the larger loss fu pre-factor. If you set we for a row, the returned prediction frame at that zero and this is incorrec an accurate prediction, all rows with weight ==
nbins_top_level	1024	For numerical columns build a histogram of (at this many bins at the ro then decrease by factor per level
nbins_cats	1024	For categorical columns (factors), build a histogr this many bins, then spl best point. Higher value lead to more overfitting

r2_stopping	1.7976931348623157e+308	r2_stopping is no longe supported and will be ig set - please use stopping_rounds, stopping_metric and stopping_tolerance inst Previous version of H2(stop making trees wher metric equals or exceec
stopping_rounds	0	Early stopping based or convergence of stopping_metric. Stop if moving average of leng the stopping_metric do improve for k:=stopping scoring events (0 to dis
stopping_metric	AUTO	Metric to use for early s (AUTO: logloss for class deviance for regression anonomaly_score for Is Forest). Note that custc custom_increasing can used in GBM and DRF Python client.
stopping_tolerance	0.001	Relative tolerance for m based stopping criterio relative improvement is least this much)
max_runtime_secs	0	Maximum allowed runt seconds for model train 0 to disable.
checkpoint		Model checkpoint to re training with.
col_sample_rate_per_tree	1	Column sample rate pe (from 0.0 to 1.0)
min_split_improvement	0.00001	Minimum relative impro in squared error reducti split to happen
histogram_type	AUTO	What type of histogram for finding optimal split
categorical_encoding	AUTO	Encoding scheme for ca features
distribution	AUTO	Distribution function
custom_metric_func		Reference to custom ev function, format: `language:keyName=fu
export_checkpoints_dir		Automatically export ge models to this directory
gainslift_bins	-1	Gains/Lift table numbe 0 means disabled.. Defa -1 means automatic bin
auc_type	AUTO	Set default multinomial type.

EXPERT

build_tree_one_node	☐	Run on one node only; network overhead but f cpus used. Suitable for datasets.

sample_rate_per_class		A list of row sample rat class (relative fraction f class, from 0.0 to 1.0), f tree
binomial_double_trees	☐	For binary classification 2x as many trees (one p - can lead to higher acc
col_sample_rate_change_per_level	1	Relative change of the sampling rate for every (must be > 0.0 and <= 2
calibrate_model	☐	Use Platt Scaling to calc calibrated class probabi Calibration can provide accurate estimates of c probabilities.
calibration_frame	(Choose...)	Calibration frame for Pl Scaling
check_constant_response	☑	Check if response colur constant. If enabled, th exception is thrown if t response column is a cc value.If disabled, then n will train regardless of t response column being constant value or not.

Build Model

Job

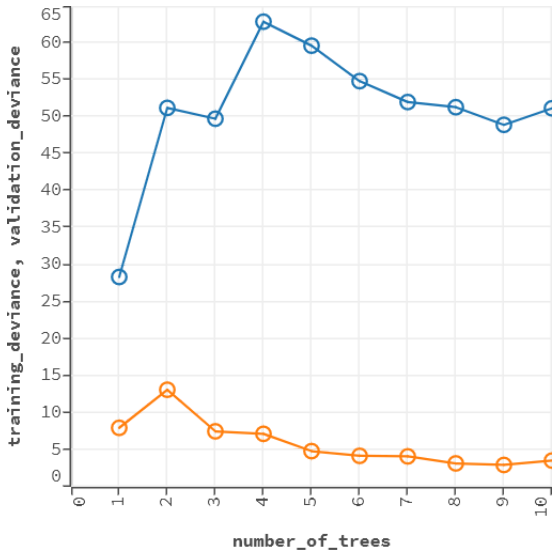
Run Time 00:00:00.119
Remaining Time 00:00:00.0
Type Model
Key drf-b5713458-ede7-46c5-b920-b5ecbea620c7
Description DRF
Status DONE
Progress 100%
Done.
Actions View

Model

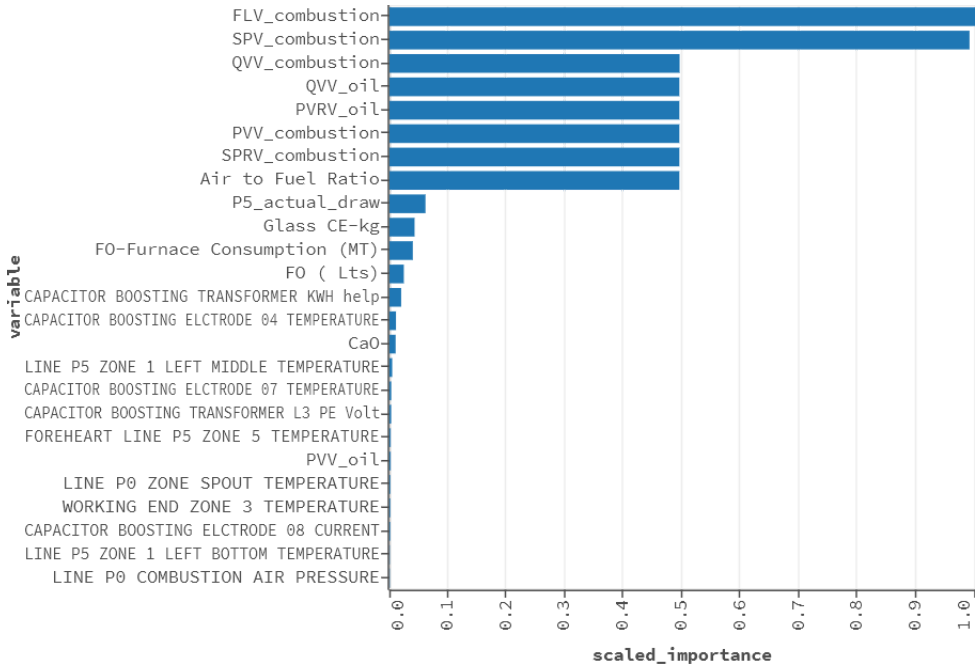
Model ID: drf-b5713458-ede7-46c5-b920-b5ecbea620c7
Algorithm: Distributed Random Forest
Actions: Refresh | Predict... | Download POJO | Download Model Deployment Package (MOJO) | Export | Inspect | Delete | Download Gen Model

▸ MODEL PARAMETERS

▾ SCORING HISTORY - DEVIANCE



▾ VARIABLE IMPORTANCES



▸ OUTPUT

▸ COLUMN_TYPES

▸ OUTPUT - MODEL SUMMARY

▸ OUTPUT - SCORING HISTORY

▸ OUTPUT - TRAINING_METRICS

▸ OUTPUT - VALIDATION_METRICS

▸ OUTPUT - VARIABLE IMPORTANCES

▾ PREVIEW POJO

Preview POJO

CS

Partial Dependence

Save Destination PDP as: pdp-6bfdfe39-ef92-407e-8425-a2fe9e

Model: drf-b5713458-ede7-46c5-b920-b5ecbea620c7

Frame:	flatfile_prod2.hex	
row_index	-1	Row for which partial dependence will be calculated instead of the whole input frame (-1 for all).
nbins	20	How many levels should PDP compute. More levels will make it slower.
Select columns?	☐	Checking this will allow you to select custom columns for PDP. By default, the top 10 features are used. Those features are sorted by variable importance.
2D PDP Columns:	+ Add	Select lists of column name pairs to plot 2D partial dependence plot for.
Actions:	Compute	

Job

Run Time	00:00:00.181
Remaining Time	00:00:00.0
Type	PartialDependence
Key	pdp-6bfdfe39-ef92-407e-8425-a2fe9e9e8aab
Description	PartialDependence
Status	DONE
Progress	100%
	Done.
Actions	View

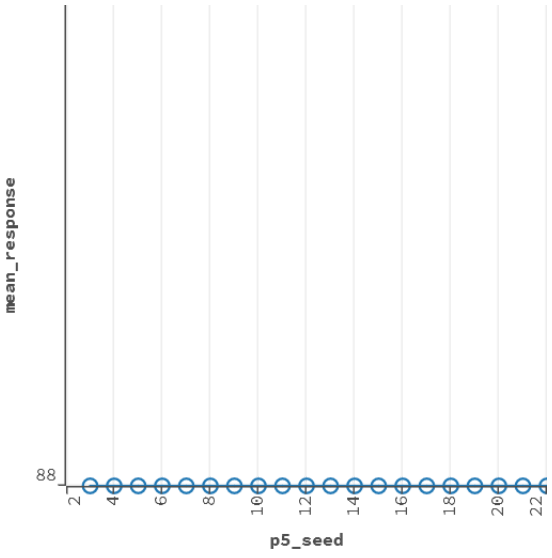
Partial Dependence Summary

Model ID: drf-b5713458-ede7-46c5-b920-b5ecbea620c7

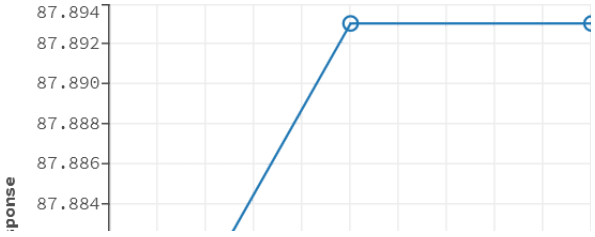
Frame ID: flatfile_prod2.hex

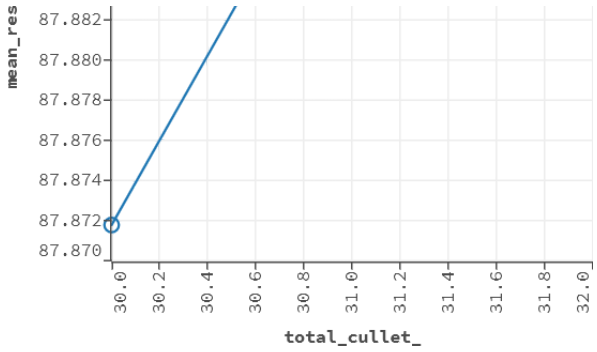
Show PDP Data Table? ☐

▾ PARTIAL DEPENDENCE PLOT OF MODEL DRF-B5713458-EDE7-46C5-B920-B5ECBEA620C7 ON COLUMN 'P5_SEED'.

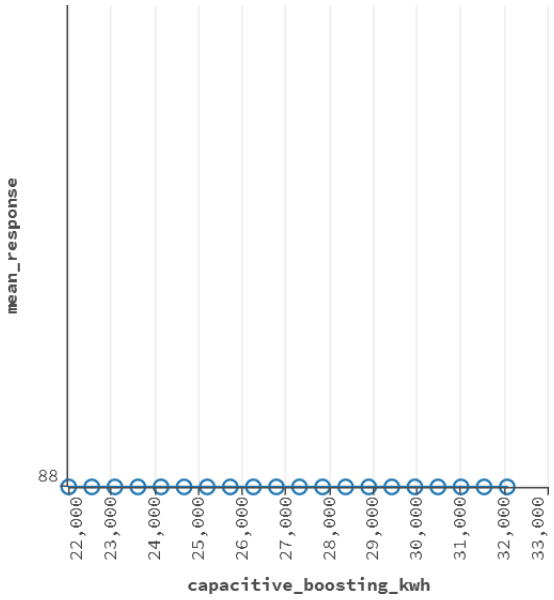


▾ PARTIAL DEPENDENCE PLOT OF MODEL DRF-B5713458-EDE7-46C5-B920-B5ECBEA620C7 ON COLUMN 'TOTAL CULLET %'.

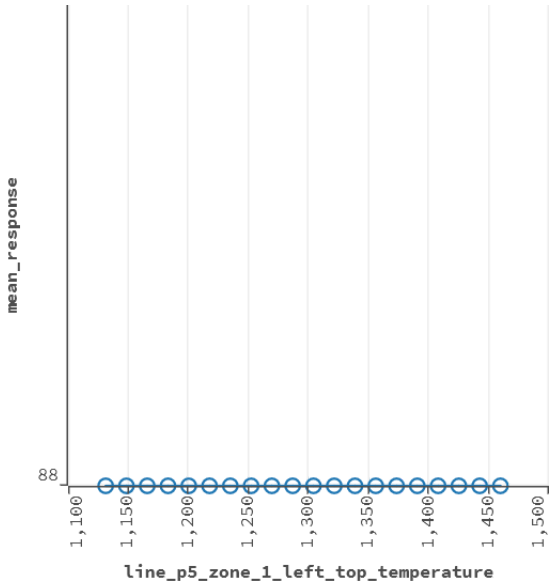




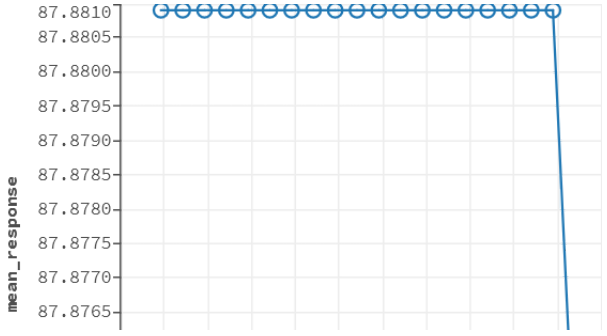
▾ PARTIAL DEPENDENCE PLOT OF MODEL DRF-B5713458-EDE7-46C5-B920-B5ECBEA620C7 ON COLUMN 'CAPACITIVE BOOSTING (KWH)'.

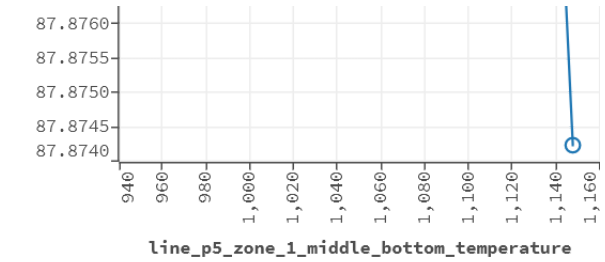


▾ PARTIAL DEPENDENCE PLOT OF MODEL DRF-B5713458-EDE7-46C5-B920-B5ECBEA620C7 ON COLUMN 'LINE P5 ZONE 1 LEFT TOP TEMPERATURE'.

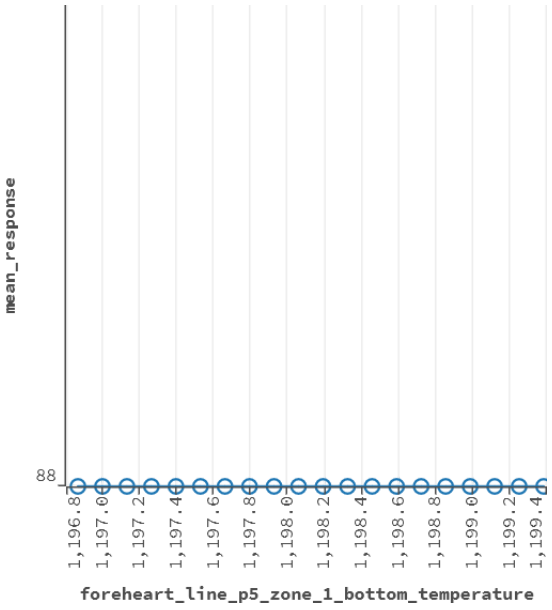


▾ PARTIAL DEPENDENCE PLOT OF MODEL DRF-B5713458-EDE7-46C5-B920-B5ECBEA620C7 ON COLUMN 'LINE P5 ZONE 1 MIDDLE BOTTOM TEMPERATURE'.

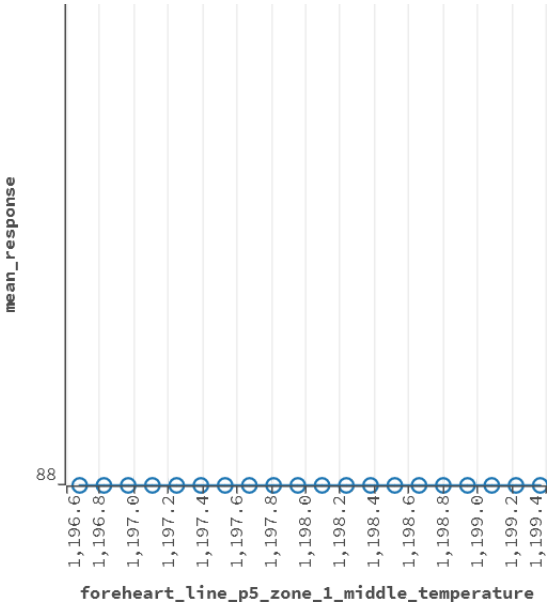




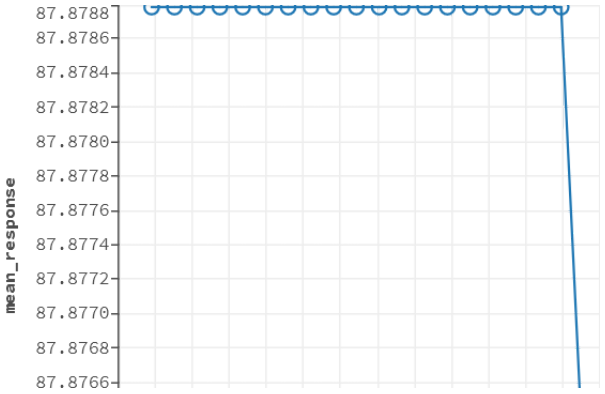
▾ PARTIAL DEPENDENCE PLOT OF MODEL DRF-B5713458-EDE7-46C5-B920-B5ECBEA620C7 ON COLUMN 'FOREHEART LINE P5 ZONE 1 BOTTOM TEMPERATURE'.

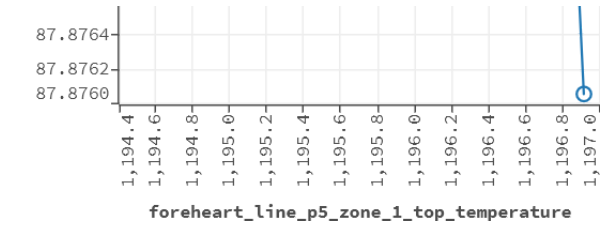


▾ PARTIAL DEPENDENCE PLOT OF MODEL DRF-B5713458-EDE7-46C5-B920-B5ECBEA620C7 ON COLUMN 'FOREHEART LINE P5 ZONE 1 MIDDLE TEMPERATURE'.

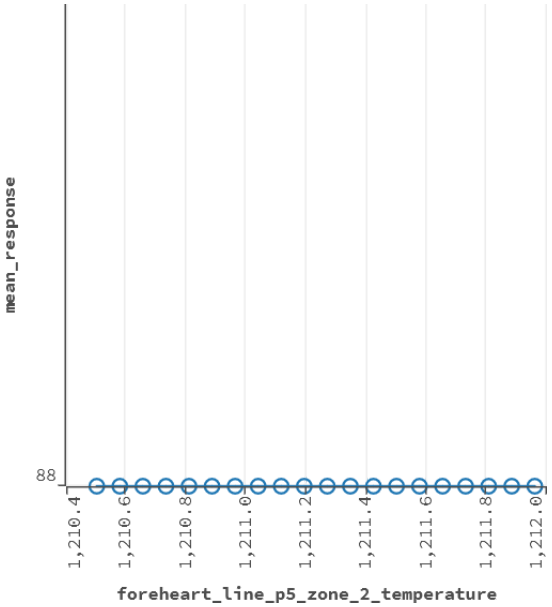


▾ PARTIAL DEPENDENCE PLOT OF MODEL DRF-B5713458-EDE7-46C5-B920-B5ECBEA620C7 ON COLUMN 'FOREHEART LINE P5 ZONE 1 TOP TEMPERATURE'.

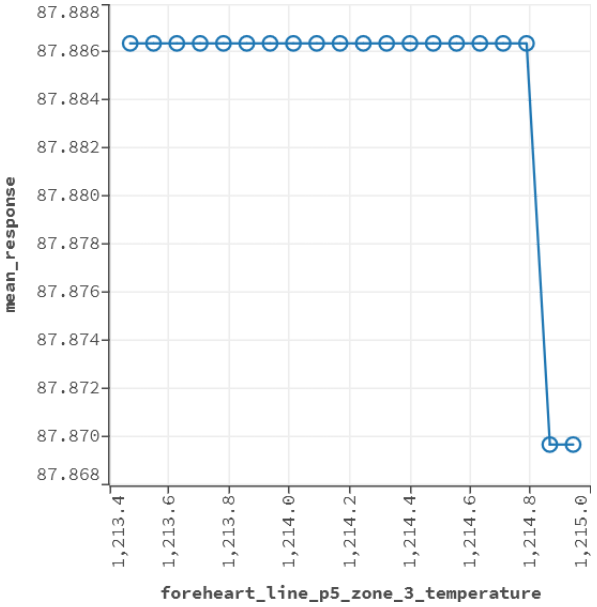




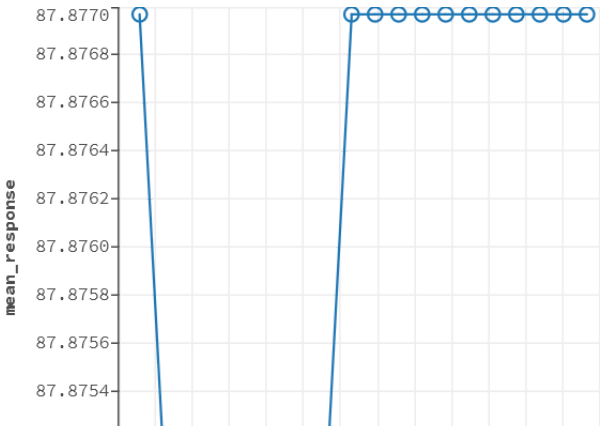
▾ PARTIAL DEPENDENCE PLOT OF MODEL DRF-B5713458-EDE7-46C5-B920-B5ECBEA620C7 ON COLUMN 'FOREHEART LINE P5 ZONE 2 TEMPERATURE'.

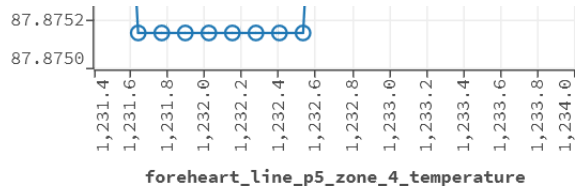


▾ PARTIAL DEPENDENCE PLOT OF MODEL DRF-B5713458-EDE7-46C5-B920-B5ECBEA620C7 ON COLUMN 'FOREHEART LINE P5 ZONE 3 TEMPERATURE'.

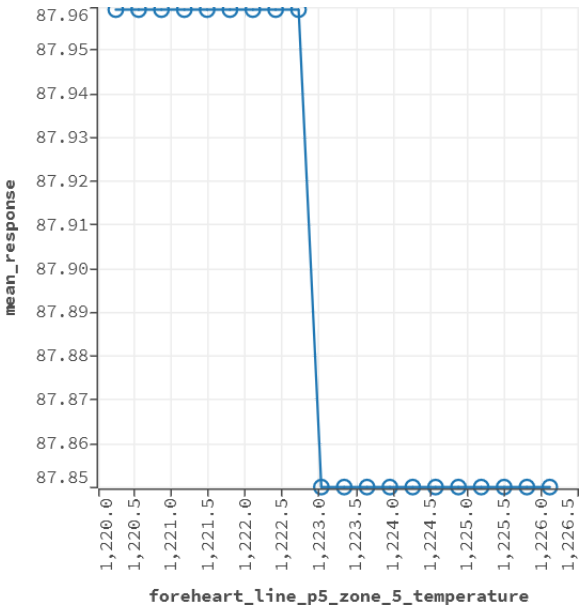


▾ PARTIAL DEPENDENCE PLOT OF MODEL DRF-B5713458-EDE7-46C5-B920-B5ECBEA620C7 ON COLUMN 'FOREHEART LINE P5 ZONE 4 TEMPERATURE'.

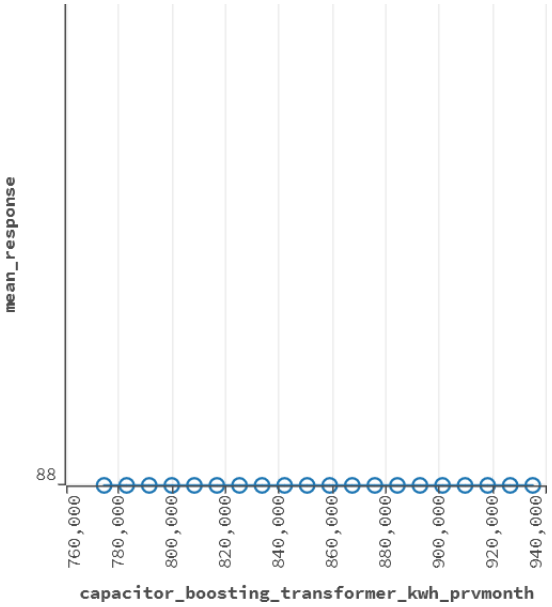




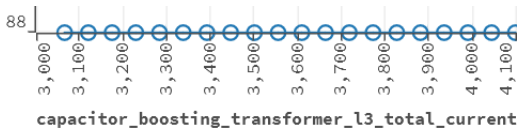
▾ PARTIAL DEPENDENCE PLOT OF MODEL DRF-B5713458-EDE7-46C5-B920-B5ECBEA620C7 ON COLUMN 'FOREHEART LINE P5 ZONE 5 TEMPERATURE'.



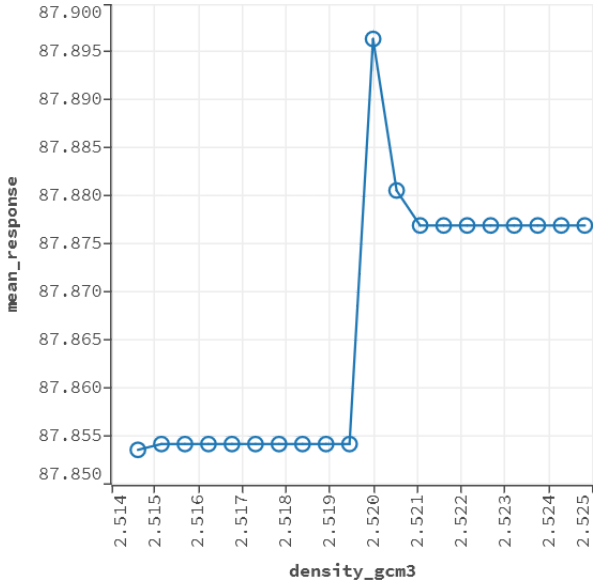
▾ PARTIAL DEPENDENCE PLOT OF MODEL DRF-B5713458-EDE7-46C5-B920-B5ECBEA620C7 ON COLUMN 'CAPACITOR BOOSTING TRANSFORMER KWH PRVMONTH'.



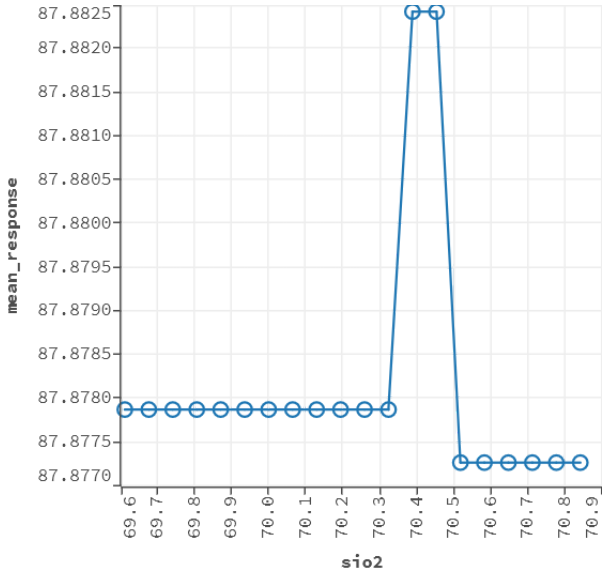
▾ PARTIAL DEPENDENCE PLOT OF MODEL DRF-B5713458-EDE7-46C5-B920-B5ECBEA620C7 ON COLUMN 'CAPACITOR BOOSTING TRANSFORMER L3 TOTAL CURRENT'.



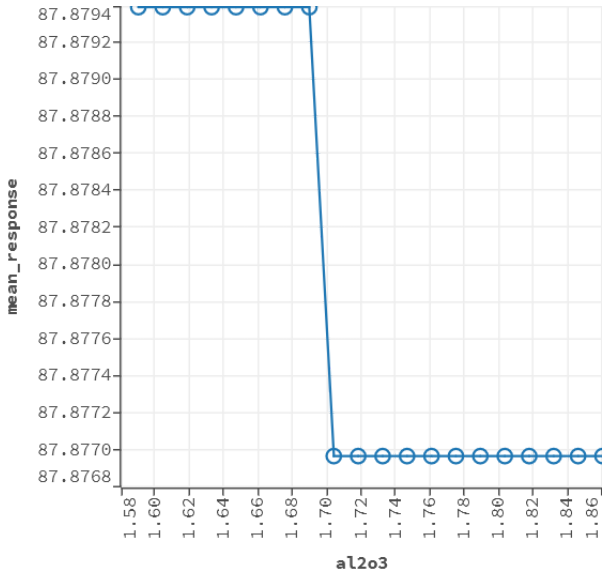
▾ PARTIAL DEPENDENCE PLOT OF MODEL DRF-B5713458-EDE7-46C5-B920-B5ECBEA620C7 ON COLUMN 'DENSITY (G/CM3)'.



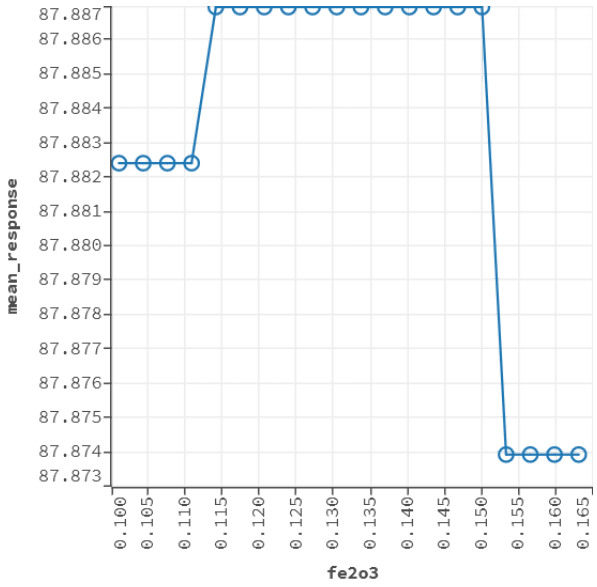
▾ PARTIAL DEPENDENCE PLOT OF MODEL DRF-B5713458-EDE7-46C5-B920-B5ECBEA620C7 ON COLUMN 'SIO2'.



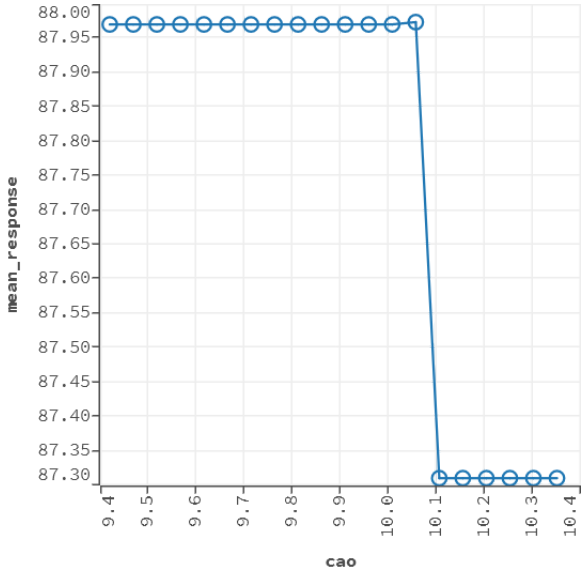
▾ PARTIAL DEPENDENCE PLOT OF MODEL DRF-B5713458-EDE7-46C5-B920-B5ECBEA620C7 ON COLUMN 'AL2O3'.



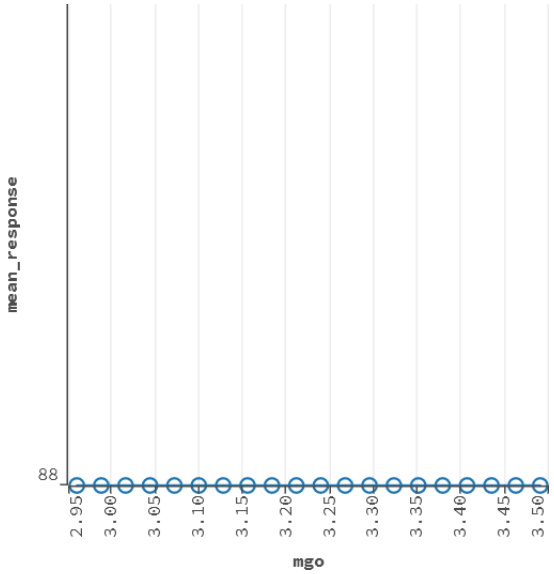
▾ PARTIAL DEPENDENCE PLOT OF MODEL DRF-B5713458-EDE7-46C5-B920-B5ECBEA620C7 ON COLUMN 'FE2O3'.



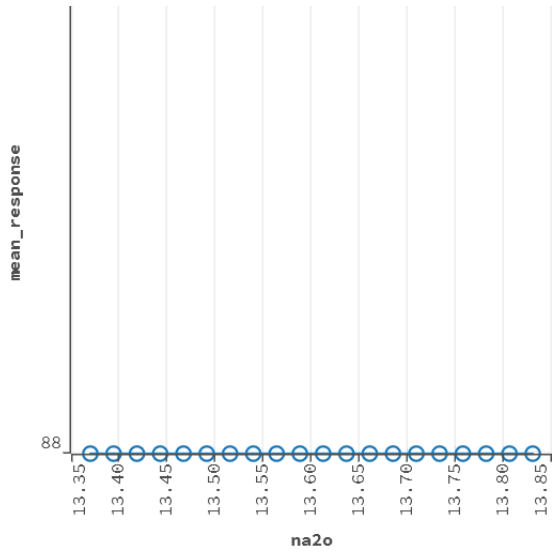
▾ PARTIAL DEPENDENCE PLOT OF MODEL DRF-B5713458-EDE7-46C5-B920-B5ECBEA620C7 ON COLUMN 'CAO'.



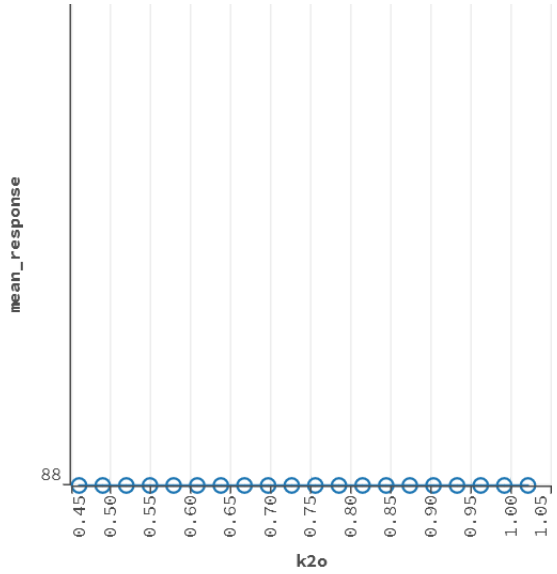
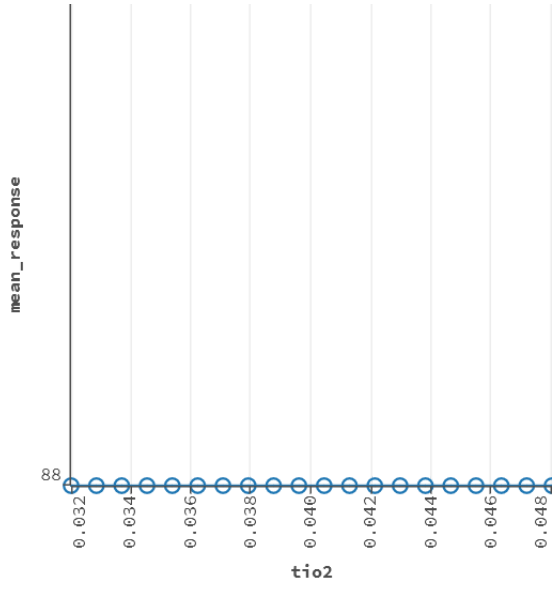
▾ PARTIAL DEPENDENCE PLOT OF MODEL DRF-B5713458-EDE7-46C5-B920-B5ECBEA620C7 ON COLUMN 'MGO'.



▾ PARTIAL DEPENDENCE PLOT OF MODEL DRF-B5713458-EDE7-46C5-B920-B5ECBEA620C7 ON COLUMN 'NA2O'.



▾ PARTIAL DEPENDENCE PLOT OF MODEL DRF-B5713458-EDE7-46C5-B920-B5ECBEA620C7 ON COLUMN 'K2O'.



▾ PARTIAL DEPENDENCE PLOT OF MODEL DRF-B5713458-EDE7-46C5-B920-B5ECBEA620C7 ON COLUMN 'TIO2'.