














PGP SL Eff Prod 1

Assistance

Routine	Description
 importFiles	Import file(s) into H ₂ O
 importSqlTable	Import SQL table into H ₂ O
 getFrames	Get a list of frames in H ₂ O
 splitFrame	Split a frame into two or more frames
 mergeFrames	Merge two frames into one
 getModels	Get a list of models in H ₂ O
 getGrids	Get a list of grid search results in H ₂ O
 getPredictions	Get a list of predictions in H ₂ O
 getJobs	Get a list of jobs running in H ₂ O
 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\flat file 180ML DCSL LW HIP - o



Search Results: (All files added)

Selected Files: 1 file selected: Clear All

✖ C:\Users\ritwi\Desktop\ripik ai\Sri Lanka Efficiency Use Case\flat file 180ML DCSL LW HIP - outlier removed.csv


Actions:

 Import

 1 / 1 files imported.


Files  C:\Users\ritwi\Desktop\ripik ai\Sri Lanka Efficiency Use Case\flat file 180ML DCSL LW HIP - outlier removed.csv

Actions

 Parse these files...

Setup Parse

PARSE CONFIGURATION

Sources  nfs:C:\Users\ritwi\Desktop\ripik ai\Sri Lanka Efficiency Use Case\flat file 180ML DCSL LW HIP - outlier removed.csv

ID flat_file_180ML_DCSL_LW_HIP___outlier_removed.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

Search by column name...

1	Date	Time	16-Dec-21	17-Dec-21	18-Dec-21	19-Dec-21	06-Jan-22	07-Jan-22
2	Adj_eff	Numeric	86.78770036	85.3426893	88.30009807	81.12080103	84.57968192	88.35278426
3	P5_eff	Numeric	91.3	92	91.9	91.8	81.1	83.3
4	P5_sap_dra	Numeric	56.78	57.2	56.81	56.95	56.37	56.19
5	P5_actual_	Numeric	56.8	57.2	56.8	57	56.4	56.2
6	P5_seed	Numeric	14	22	24	21	20	7
7	Total SAP I	Numeric	305.99	306.74	306.16	309.86	308.68	302.49
8	Adjusted D	Numeric	305.7	306.4	305.9	309.6	308.1	299.9
9	Total Cull	Numeric	30	30	30	30	31	31
10	Melting co	Numeric	10393	10285	10375	10359	10383	10617
11	Barrier Bo	Numeric	5989	5808	5562	6346	7706	6838
12	Capacitive	Numeric	30410	30526	31083	31047	31997	30786
13	Boosting (I	Numeric	36399	36334	36645	37393	39703	37624
14	Barrier Bo	Numeric	19.57	18.93	18.17	20.48	24.96	22.61
15	Capacitive	Numeric	99.38	99.52	101.53	100.2	103.66	101.78

← Previous page

→ Next page

Parse

Job

Run Time

00:00:00.36

Remaining Time

00:00:00.0

Type

Frame

Key

flat_file_180ML_DCSL_LW_HIP___outlier_removed.hex

Description

Parse

Status

DONE

Progress

100%

Done.

Actions

View

flat_file_180ML_DCSL_LW_HIP___outlier_removed.hex

Actions:

View Data

Split

Build Model

Run AutoML

Predict

Download

Delete

Export

Rows	Columns	Compressed Size
35	157	33KB

COLUMN SUMMARIES

Barrier Boosting (kWh)	int	0	0	0	0	3922.0	8312.0	6586.1429	103
Capacitive Boosting (kWh)	int	0	0	0	0	26403.0	31997.0	30741.4857	90
Boosting (kWh)	int	0	0	0	0	31333.0	39703.0	37327.6286	171
Barrier Boosting Per ton (kWh/ton)	real	0	0	0	0	12.6000	26.7700	21.4617	103
Capacitive Boosting Per ton (kWh/ton)	real	0	0	0	0	85.3800	105.4500	100.1363	90

F0-Furnace Consumption (MT)	type	Missing	Zeros	+Inf	-Inf	min	max	mean
	real	0	0	0	0	27.2310	28.9280	28.3586
F0-Furnace Consumption (kg per MT)	real	0	0	0	0	90.4700	95.2900	92.3640
LPG -WE/FHs Consumption(kg)	int	0	0	0	0	5212.0	5874.0	5527.4571
Melting Energy -Kcal//Draw of glass	int	0	0	0	0	985.0	1037.0	1006.4000
Bottom Front (oC)	int	0	0	0	0	1212.0	1231.0	1221.7143

← Previous 20 Columns

→ Next 20 Columns

- CHUNK COMPRESSION SUMMARY
- FRAME DISTRIBUTION SUMMARY

✂ Split Frame

Frame: flat_file_180ML_DCSL_LW_HIP___outlier_removed.hex ▾

Splits:	Ratio	Key
	0.75	frame_0.750
	0.250	frame_0.250

Add a new split

Seed: 295443

✂ Create

📊 Split Frames

Type	Key	Ratio
📊	frame_0.750	0.75
📊	frame_0.250	0.25

🏗 Build a Model

Select an algorithm: Distributed Random Forest ▾

PARAMETERS

model_id	drf-eb12b88c-f929-443f-9e22-aadee8ef0390	Destination id for this n
training_frame	frame_0.750 ▾	auto-generated if not sp
validation_frame	frame_0.250 ▾	Id of the training data fi
nfolds	0	Id of the validation dat
response_column	P5_eff ▾	Number of folds for K-f
		cross-validation (0 to di
		>= 2).
		Response variable colou

ignored_columns

Search...

Names of columns to ignore during training.

Showing page 1 of 16. -154 ignored.

<input checked="" type="checkbox"/> Date	TIME
<input checked="" type="checkbox"/> Adj_eff	REAL
<input checked="" type="checkbox"/> P5_eff	REAL
<input type="checkbox"/> P5_sap_draw	REAL
<input type="checkbox"/> P5_actual_draw	REAL
<input type="checkbox"/> P5_seed	INT
<input type="checkbox"/> Total SAP Draw	REAL
<input type="checkbox"/> Adjusted Draw	REAL
<input type="checkbox"/> Total Cullet %	INT
<input type="checkbox"/> Melting cost-Rs/ton/Draw	INT

☒ All

☐ None

← Previous 10

→ Next 10

Only show columns with more than 0 % missing values.

ignore_const_cols

☒

Ignore constant columns

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 split best point

seed -1

Seed for pseudo random number generator (if applicable)

mtries -1

Number of variables randomly sampled as candidates at each split. If set to -1, default sqrt(p) for classification or regression (where p is number of predictors

sample_rate 0.632

Row sample rate per tree (0.0 to 1.0)

ADVANCED

score_each_iteration ☐

Whether to score during iteration of model training

score_tree_interval 0

Score the model after every many trees. Disabled if score_each_iteration is enabled.

fold_column (Choose...)

▼

Column with cross-validation fold index assignment per observation.

offset_column (Choose...)

▼

Offset column. This will be added to the combination of columns before applying the function.

<i>weights_column</i>	(Choose...)	Column with observation weights. Giving some observation a weight of equivalent to excluding the dataset; giving an observation a relative weight of 2 is equivalent to repeating the row twice. Negative weights are not allowed. Note: Weights are per-row observation weights and do not increase the size of the data frame. This is the number of times a row is repeated, but non-integer values are supported as well. During training, rows with higher weights matter more due to the larger loss function pre-factor. If you set weights for a row, the returned prediction frame at that row is zero and this is incorrect for an accurate prediction, all rows with weight == 1.
<i>nbins_top_level</i>	1024	For numerical columns, build a histogram of (at this many bins at the root) then decrease by factor of 2 per level.
<i>nbins_cats</i>	1024	For categorical columns (factors), build a histogram of this many bins, then split at the best point. Higher values lead to more overfitting.
<i>r2_stopping</i>	1.7976931348623157e+308	<i>r2_stopping</i> is no longer supported and will be ignored - please use <i>stopping_rounds</i> , <i>stopping_metric</i> and <i>stopping_tolerance</i> instead. Previous version of H2O stopped making trees when <i>metric</i> equals or exceeds <i>r2_stopping</i> .
<i>stopping_rounds</i>	0	Early stopping based on convergence of <i>stopping_metric</i> . Stop if the moving average of length <i>stopping_rounds</i> of <i>stopping_metric</i> does not improve for <i>k:=stopping_rounds</i> consecutive scoring events (0 to disable).
<i>stopping_metric</i>	AUTO	Metric to use for early stopping (AUTO: logloss for classification, deviance for regression, anomaly_score for Isolation Forest). Note that custom metrics can be used in GBM and DRF via the Python client.
<i>stopping_tolerance</i>	0.001	Relative tolerance for non-based stopping criterion (relative improvement is at least this much).

<i>max_runtime_secs</i>	0	Maximum allowed run seconds for model train 0 to disable.
<i>checkpoint</i>		Model checkpoint to re training with.
<i>col_sample_rate_per_tree</i>	1	Column sample rate per (from 0.0 to 1.0)
<i>min_split_improvement</i>	0.00001	Minimum relative impr in squared error reducti split to happen
<i>histogram_type</i>	AUTO ▼	What type of histogram for finding optimal split
<i>categorical_encoding</i>	AUTO ▼	Encoding scheme for ca features
<i>distribution</i>	AUTO ▼	Distribution function
<i>custom_metric_func</i>		Reference to custom ev function, format: `language:keyName=fu
<i>export_checkpoints_dir</i>		Automatically export ge models to this directory
<i>gainslift_bins</i>	-1	Gains/Lift table number 0 means disabled.. Defa -1 means automatic bin
<i>auc_type</i>	AUTO ▼	Set default multinomial type.

EXPERT

build_tree_one_node

☐

sample_rate_per_class

binomial_double_trees

☐

col_sample_rate_change_per_level

1

calibrate_model

☐

calibration_frame

(Choose...) ▼

check_constant_response

☒

Run on one node only;
network overhead but 1
cpus used. Suitable for
datasets.

A list of row sample rat
class (relative fraction f
class, from 0.0 to 1.0), f
tree

For binary classification
2x as many trees (one p
- can lead to higher acc

Relative change of the c
sampling rate for every
(must be > 0.0 and <= 2

Use Platt Scaling to cak
calibrated class probabi
Calibration can provide
accurate estimates of c
probabilities.

Calibration frame for Pl
Scaling

Check if response colour
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 00:00:00.133

Remaining Time 00:00:00.0

Type Model

Key  drf-eb12b88c-f929-443f-9e22-aadee8ef0390

Description DRF

Status DONE

Progress 100%









Done.

Actions  View

Model

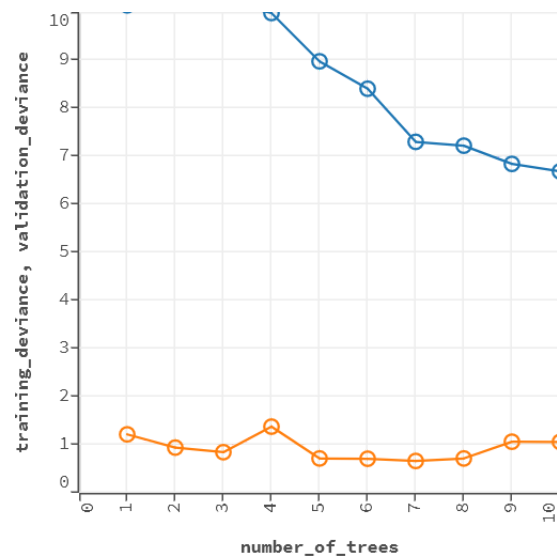
Model ID: drf-eb12b88c-f929-443f-9e22-aadee8ef0390

Algorithm: Distributed Random Forest

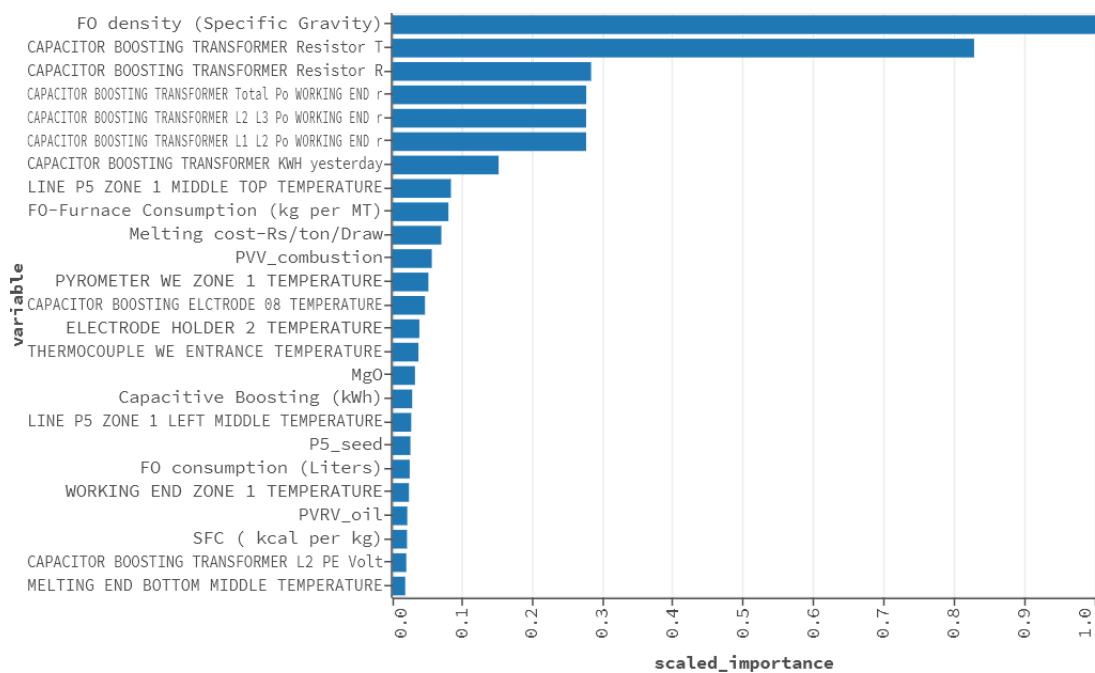
Actions:  Refresh  Predict...  Download POJO  Download Model Deployment Package (MOJO)  Export
 Inspect  Delete  Download Gen Model

▶ 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

Partial Dependence

Save Destination PDP as:

Model: ▼

Frame: ▼

row_index

Row for which partial dependence will be calculated instead of the whole input frame (-1 for all).

nbins

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.

Available columns: Search...

Showing page 2 of 2. 23 selected for PDP calculations.

<input type="checkbox"/>	CAPACITOR BOOSTING TRANSFORMER BM POWER SPH	INT
<input type="checkbox"/>	CAPACITOR BOOSTING TRANSFORMER BM POWER SPPROD	INT
<input type="checkbox"/>	CAPACITOR BOOSTING TRANSFORMER BM PRODUCTION	INT
<input type="checkbox"/>	CAPACITOR BOOSTING TRANSFORMER KWH help	REAL
<input type="checkbox"/>	CAPACITOR BOOSTING TRANSFORMER KWH memo	INT
<input type="checkbox"/>	CAPACITOR BOOSTING TRANSFORMER KWH presmonth	REAL
<input type="checkbox"/>	CAPACITOR BOOSTING TRANSFORMER KWH prvmmonth	INT
<input type="checkbox"/>	CAPACITOR BOOSTING TRANSFORMER KWH today	REAL
<input type="checkbox"/>	CAPACITOR BOOSTING TRANSFORMER KWH yesterday	REAL
<input type="checkbox"/>	CAPACITOR BOOSTING TRANSFORMER KWH today	REAL

☒ All ☐ None

Previous 100Next 100

2D PDP Columns: + Add

Select lists of column name pairs to plot 2D partial dependence plot for.

Actions: Compute

Job

Run Time

00:00:00.135

Remaining Time

00:00:00.0

Type

PartialDependence

Key

pdp-b5496bf8-1a11-41c9-b28f-a0ed2ee61cf3

Description

PartialDependence

Status

DONE

Progress

100%

Done.

Actions

View

Partial Dependence Summary

Model ID:

drf-eb12b88c-f929-443f-9e22-aadee8ef0390

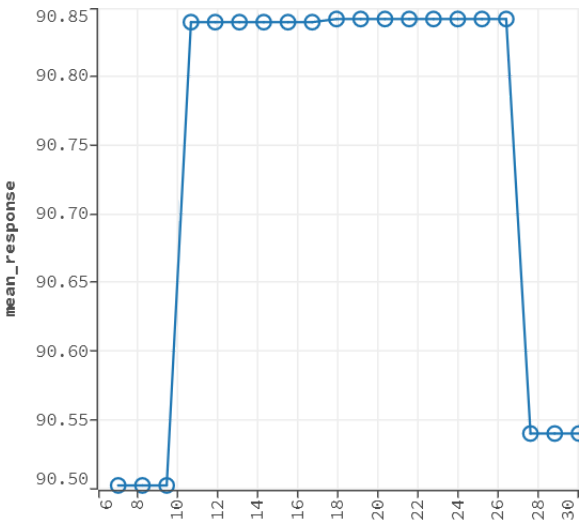
Frame ID:

flat_file_180ML_DCSL_LW_HIP___outlier_removed.hex

Show PDP Data Table?

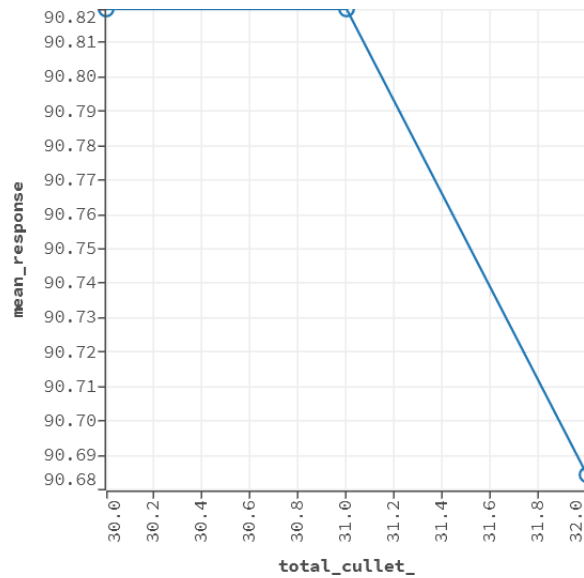
☐

PARTIAL DEPENDENCE PLOT OF MODEL DRF-EB12B88C-F929-443F-9E22-AADEE8EF0390 ON COLUMN 'P5_SEED'.

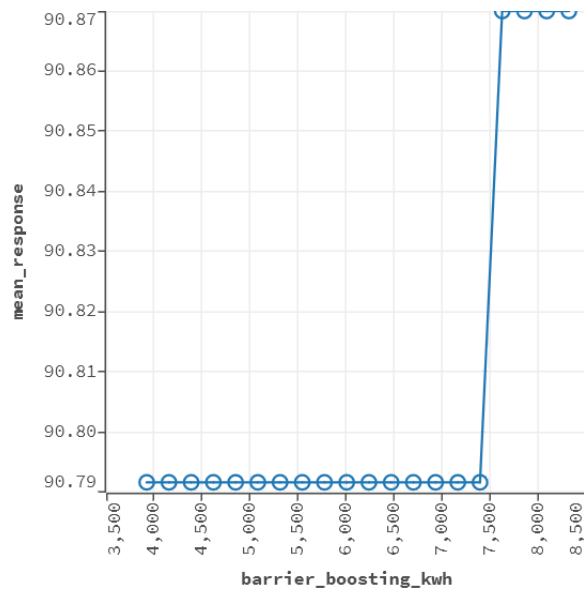


p5_seed

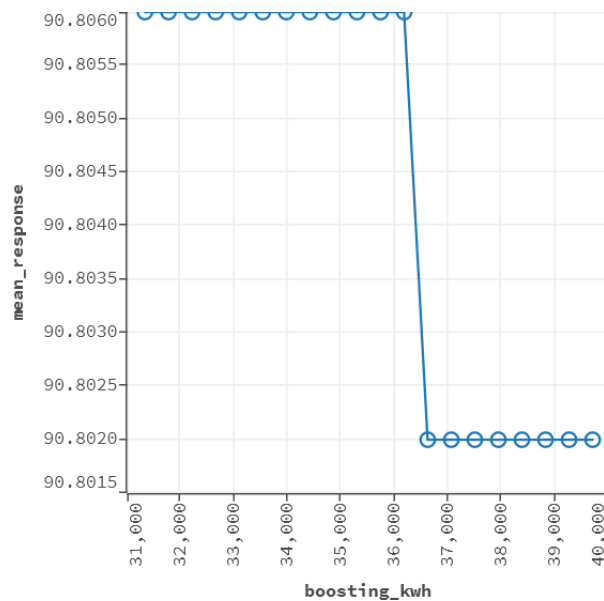
▼ PARTIAL DEPENDENCE PLOT OF MODEL DRF-EB12B88C-F929-443F-9E22-AADEE8EF0390 ON COLUMN 'TOTAL CULLET %'.



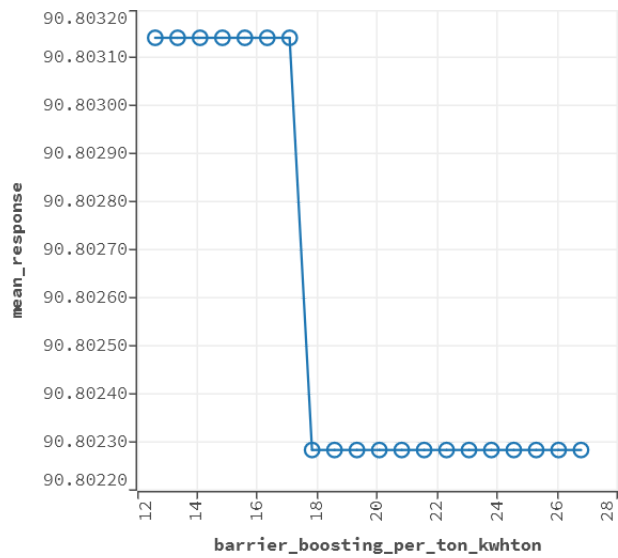
▼ PARTIAL DEPENDENCE PLOT OF MODEL DRF-EB12B88C-F929-443F-9E22-AADEE8EF0390 ON COLUMN 'BARRIER BOOSTING (KWH)'.



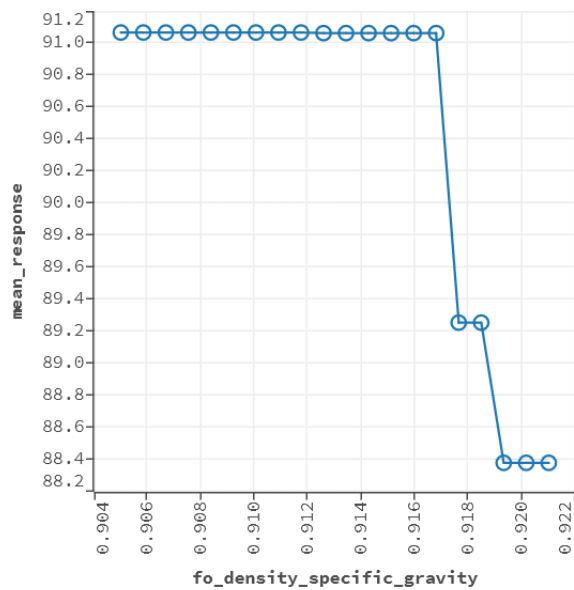
▼ PARTIAL DEPENDENCE PLOT OF MODEL DRF-EB12B88C-F929-443F-9E22-AADEE8EF0390 ON COLUMN 'BOOSTING (KWH)'.



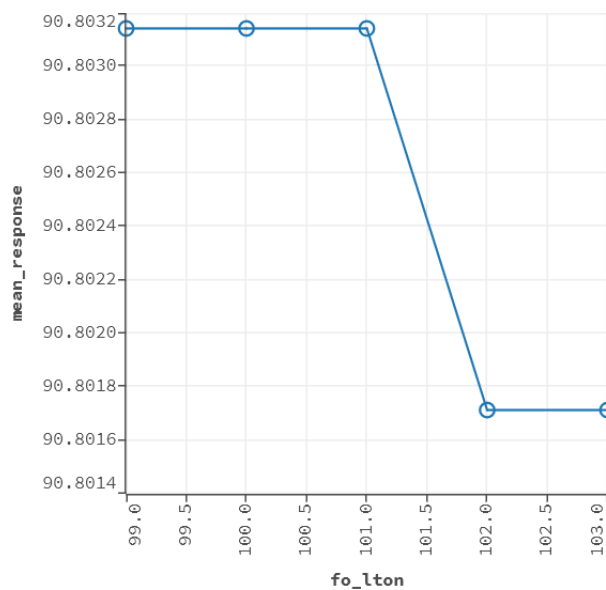
▼ PARTIAL DEPENDENCE PLOT OF MODEL DRF-EB12B88C-F929-443F-9E22-AADEE8EF0390 ON COLUMN 'BARRIER BOOSTING PER TON (KWH/TON)'.



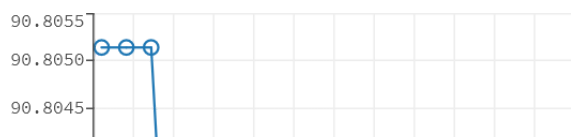
▼ PARTIAL DEPENDENCE PLOT OF MODEL DRF-EB12B88C-F929-443F-9E22-AADEE8EF0390 ON COLUMN 'FO DENSITY (SPECIFIC GRAVITY)'.

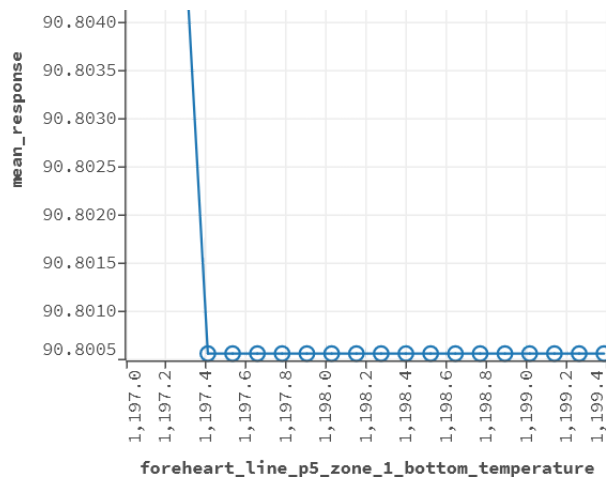


▼ PARTIAL DEPENDENCE PLOT OF MODEL DRF-EB12B88C-F929-443F-9E22-AADEE8EF0390 ON COLUMN 'FO L/TON'.

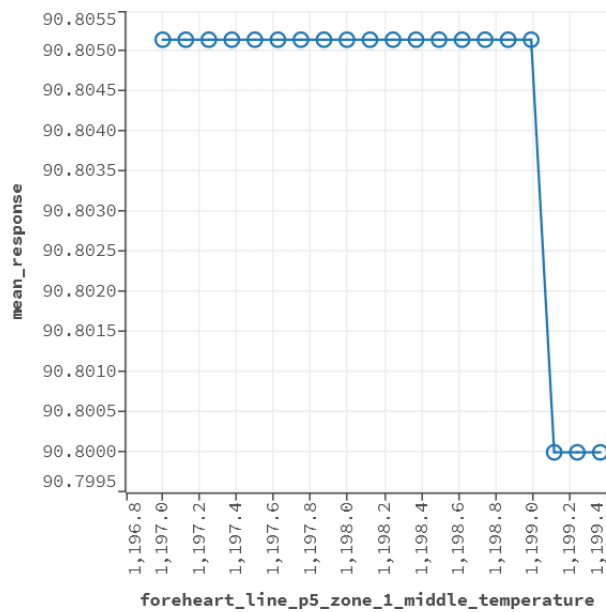


▼ PARTIAL DEPENDENCE PLOT OF MODEL DRF-EB12B88C-F929-443F-9E22-AADEE8EF0390 ON COLUMN 'FOREHEART LINE P5 ZONE 1 BOTTOM TEMPERATURE'.

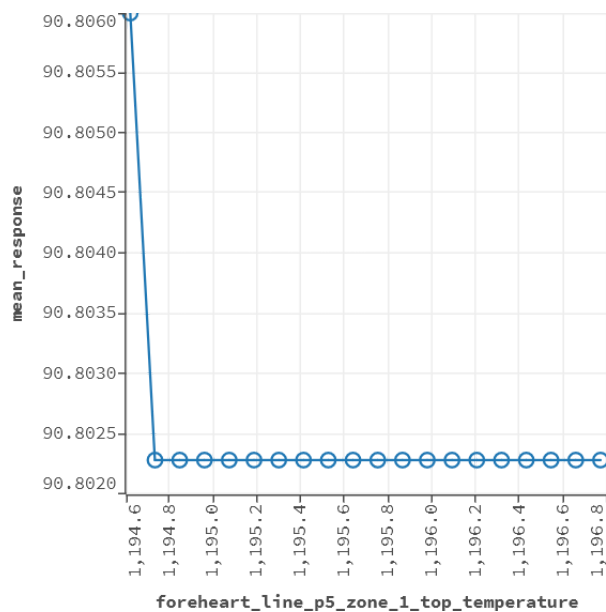




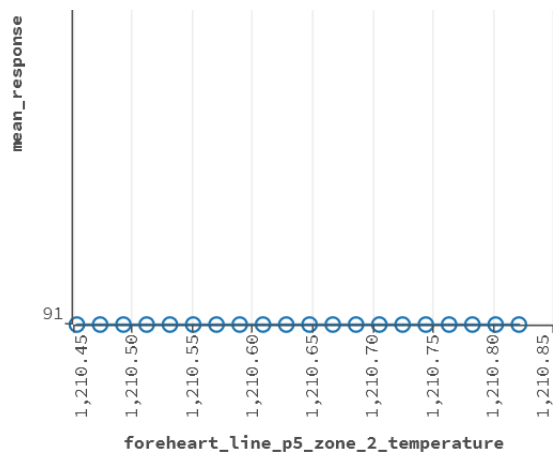
▼ PARTIAL DEPENDENCE PLOT OF MODEL DRF-EB12B88C-F929-443F-9E22-AADEE8EF0390 ON COLUMN 'FOREHEART LINE P5 ZONE 1 MIDDLE TEMPERATURE'.



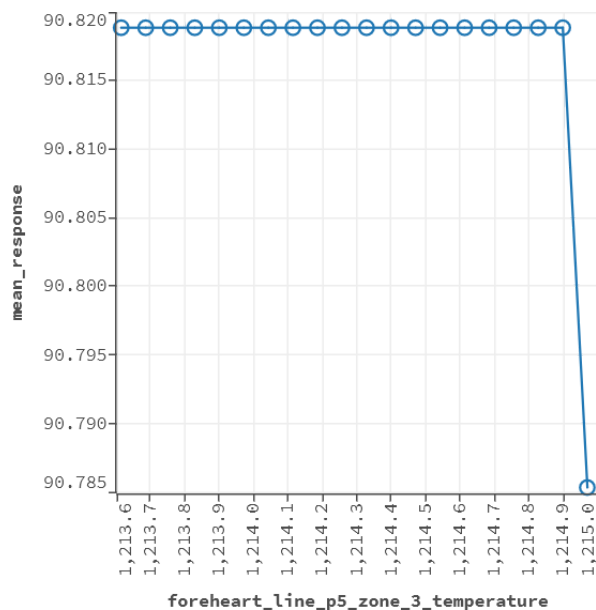
▼ PARTIAL DEPENDENCE PLOT OF MODEL DRF-EB12B88C-F929-443F-9E22-AADEE8EF0390 ON COLUMN 'FOREHEART LINE P5 ZONE 1 TOP TEMPERATURE'.



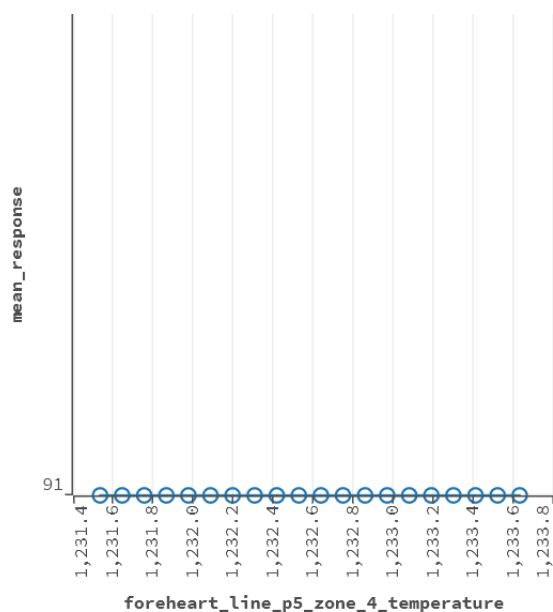
▼ PARTIAL DEPENDENCE PLOT OF MODEL DRF-EB12B88C-F929-443F-9E22-AADEE8EF0390 ON COLUMN 'FOREHEART LINE P5 ZONE 2 TEMPERATURE'.



▼ PARTIAL DEPENDENCE PLOT OF MODEL DRF-EB12B88C-F929-443F-9E22-AADEE8EF0390 ON COLUMN 'FOREHEART LINE P5 ZONE 3 TEMPERATURE'.

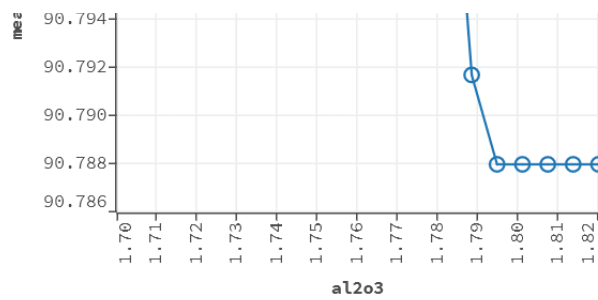


▼ PARTIAL DEPENDENCE PLOT OF MODEL DRF-EB12B88C-F929-443F-9E22-AADEE8EF0390 ON COLUMN 'FOREHEART LINE P5 ZONE 4 TEMPERATURE'.

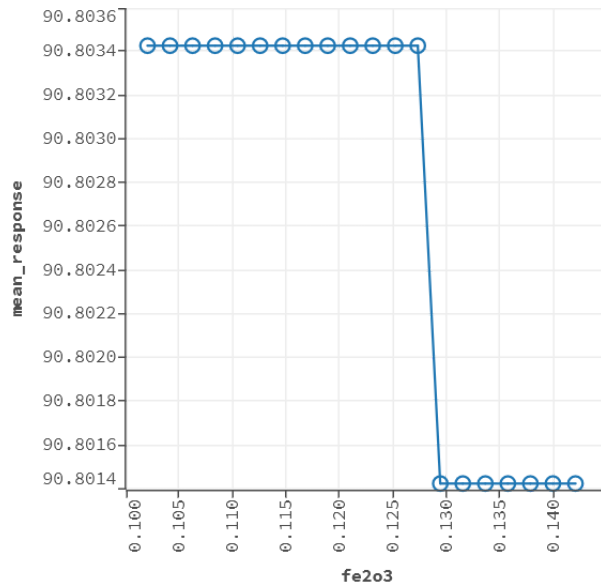


▼ PARTIAL DEPENDENCE PLOT OF MODEL DRF-EB12B88C-F929-443F-9E22-AADEE8EF0390 ON COLUMN 'FOREHEART LINE P5 ZONE 5 TEMPERATURE'.

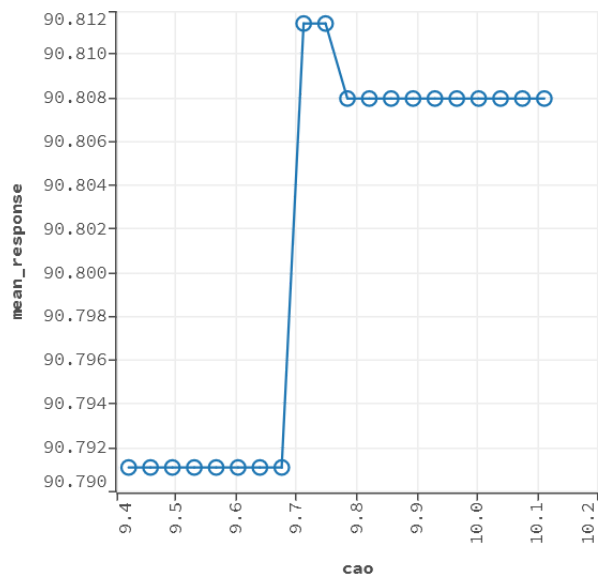




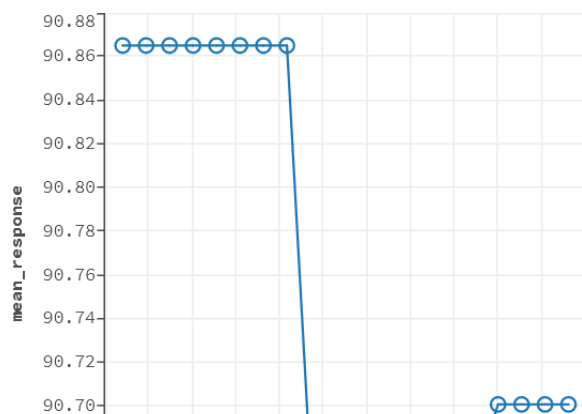
▼ PARTIAL DEPENDENCE PLOT OF MODEL DRF-EB12B88C-F929-443F-9E22-AADEE8EF0390 ON COLUMN 'FE2O3'.

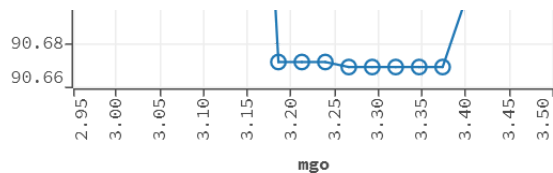


▼ PARTIAL DEPENDENCE PLOT OF MODEL DRF-EB12B88C-F929-443F-9E22-AADEE8EF0390 ON COLUMN 'CAO'.

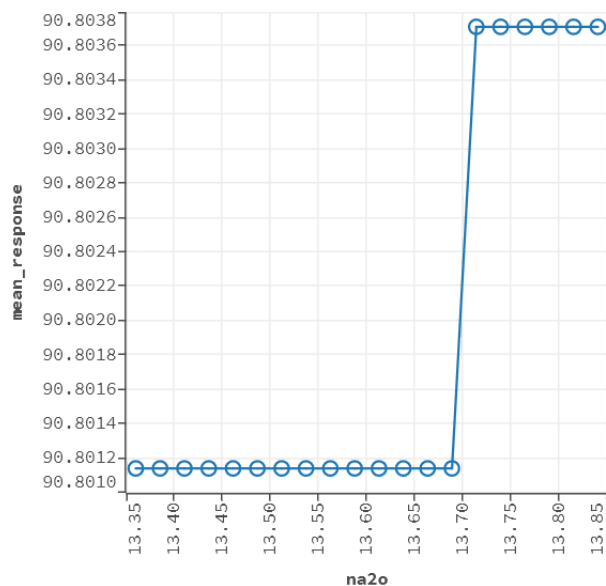


▼ PARTIAL DEPENDENCE PLOT OF MODEL DRF-EB12B88C-F929-443F-9E22-AADEE8EF0390 ON COLUMN 'MGO'.

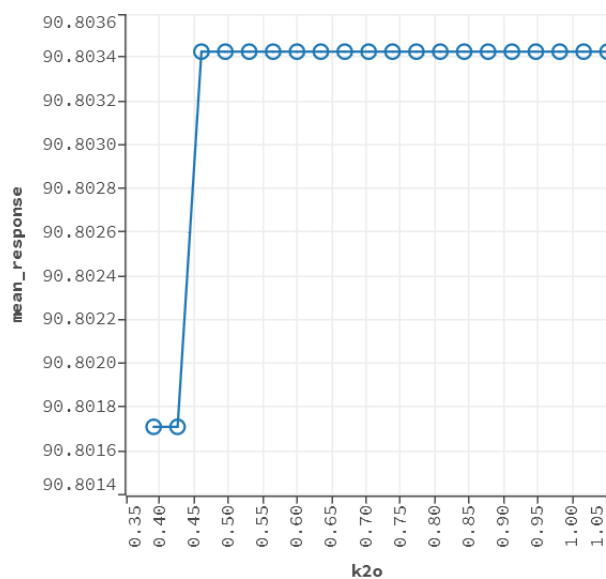




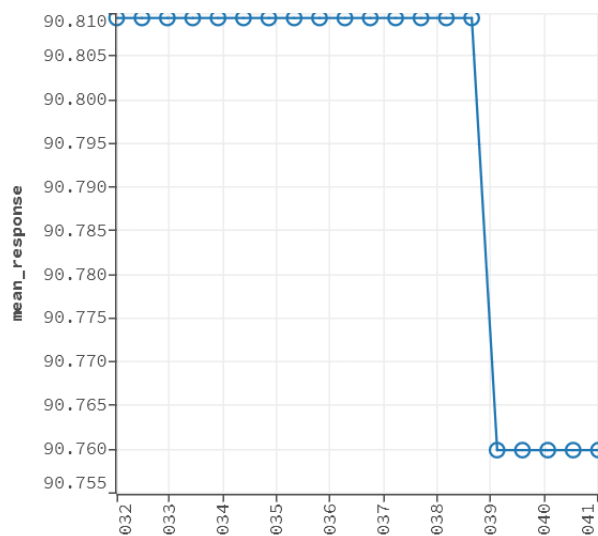
▼ PARTIAL DEPENDENCE PLOT OF MODEL DRF-EB12B88C-F929-443F-9E22-AADEE8EF0390 ON COLUMN 'NA2O'.



▼ PARTIAL DEPENDENCE PLOT OF MODEL DRF-EB12B88C-F929-443F-9E22-AADEE8EF0390 ON COLUMN 'K2O'.



▼ PARTIAL DEPENDENCE PLOT OF MODEL DRF-EB12B88C-F929-443F-9E22-AADEE8EF0390 ON COLUMN 'TiO2'.





0. 0. 0. 0. 0. 0. 0. 0. 0. 0.

tio2

