binary_classification_automl_v2

❷ 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

CS

⩫ Job

Run Time	00:36:22.481
Remaining Time	00:00:00.0
Type	Auto Model
Key	🔍 AutoML_20200209_85106537@@did_crash_happen
Description	AutoML build
Status	DONE
Progress	100%
	Done.
Actions	🔍 View

⩫ Leaderboard

🔁 Monitor Live

▼ M O D E L S

models sorted in order of aucpr, best first

	model_id	aucpr	auc
0	DRF_1_AutoML_20200209_085106	0.5317089141640107	0.78992529216
1	StackedEnsemble_BestOfFamily_AutoML_20200209_085106	0.44657386688372674	0.74558103019
2	GBM_5_AutoML_20200209_085106	0.4464852196056803	0.74420732371
3	StackedEnsemble_AllModels_AutoML_20200209_085106	0.4415597609988085	0.74215003807
4	XRT_1_AutoML_20200209_085106	0.4407658255858969	0.74549317585
5	GBM_4_AutoML_20200209_085106	0.4340007599281758	0.73654721253
6	GBM_3_AutoML_20200209_085106	0.4306078727665	0.73403632499
7	GBM_2_AutoML_20200209_085106	0.42873141272639775	0.73277032366
8	GBM_1_AutoML_20200209_085106	0.42727027933953887	0.73178809898
9	GBM_grid__1_AutoML_20200209_085106_model_1	0.420911373406698	0.72984974339
10	DeepLearning_1_AutoML_20200209_085106	0.4195079742971737	0.72644223472
11	GLM_1_AutoML_20200209_085106	0.3413556696010996	0.66615353485

▾ E V E N T   L O G

Actions taken and discoveries made by AutoML

| 40	09:15:55.420	Debug	ModelTraining	DeepLearning_1_AutoML_20200209_085106 [DeepLearn... def_1] complete
41	09:15:56.41	Info	ModelTraining	AutoML: starting XRT_1_AutoML_20200209_085106 mod... training
42	09:15:56.47	Debug	ModelTraining	XRT_1_AutoML_20200209_085106 [DRF XRT (Extremely Randomized Trees)] started
43	09:23:26.186	Debug	ModelTraining	XRT_1_AutoML_20200209_085106 [DRF XRT (Extremely Randomized Trees)] complete
44	09:23:30.310	Info	ModelTraining	AutoML: starting GBM_grid__1_AutoML_20200209_0851... hyperparameter search
45	09:23:30.316	Debug	ModelTraining	GBM_grid__1_AutoML_20200209_085106 [GBM Grid Search] started
46	09:25:47.657	Debug	ModelTraining	Built: 1 models for search: GBM_grid__1_AutoML_20200209_085106 [GBM Grid Search]
47	09:25:49.133	Debug	ModelTraining	GBM_grid__1_AutoML_20200209_085106 [GBM Grid Search] complete
48	09:25:49.133	Debug	ModelTraining	AutoML: hit the max_models limit; skipping DeepLearning grid_1
49	09:25:49.133	Debug	ModelTraining	AutoML: hit the max_models limit; skipping DeepLearning grid_2
50	09:25:49.133	Debug	ModelTraining	AutoML: hit the max_models limit; skipping DeepLearning grid_3
51	09:25:49.134	Info	ModelTraining	AutoML: starting StackedEnsemble_BestOfFamily_AutoML_20200209_0851... model training

📦 Model

Model ID: DRF_1_AutoML_20200209_085106

Algorithm: Distributed Random Forest

*Actions:*   ⟳ Refresh   ⚡ Predict...   ⬇ Download POJO

⬇ Download Model Deployment Package (MOJO)   💾 Export   ☰ Inspect   🗑 Delete

⬇ Download Gen Model

▸ MODEL PARAMETERS

▾ SCORING HISTORY - LOGLOSS
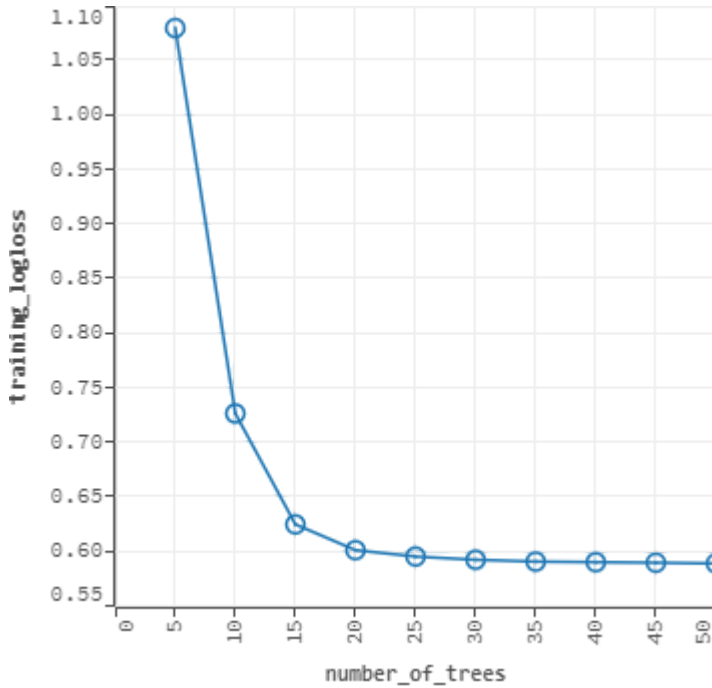
▾ ROC CURVE - TRAINING METRICS , AUC = 0.869877

Threshold:

| Choose... ▼ |

Criterion:

| max f1 ▼ |
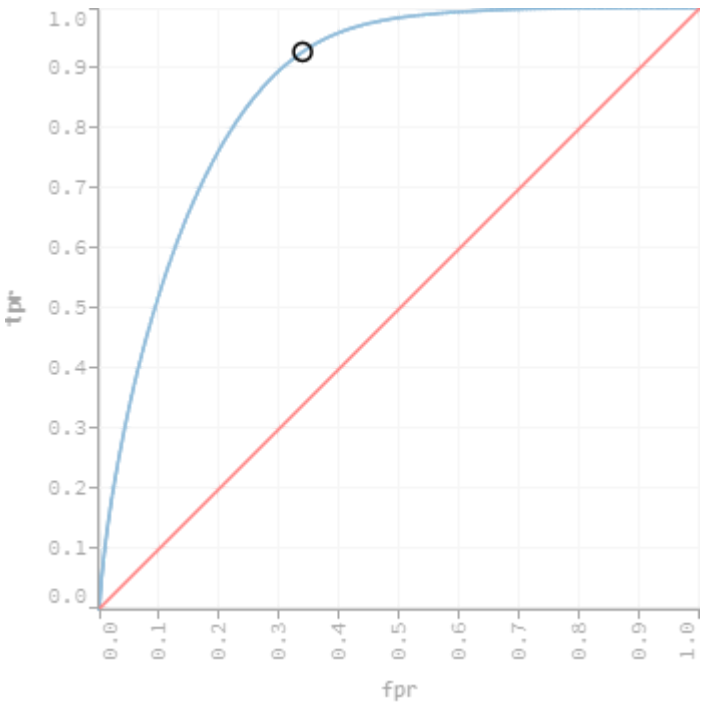
Selected mark(s):

threshold	0.2092
f1	0.8183
f2	0.8802
f0point5	0.7645
accuracy	0.7941
precision	0.7324
recall	0.9269
specificity	0.6614
absolute_mcc	0.6102
min_per_class_accuracy	0.6614
mean_per_class_accuracy	0.7941
tns	916091
fns	101244
fps	469081
tps	1284020
tnr	0.6614
fnr	0.0731
fpr	0.3386
tpr	0.9269
idx	281

	Actual/Predic
	0
CM	
	1
	Total

▾ R O C   C U R V E - C R O S S   V A L I D A T I O N   M E T R I C S ,   A U C   =   0 . 7 0 8 4 9 6

Threshold:

Choose... ▼

Criterion:

max f1 ▼
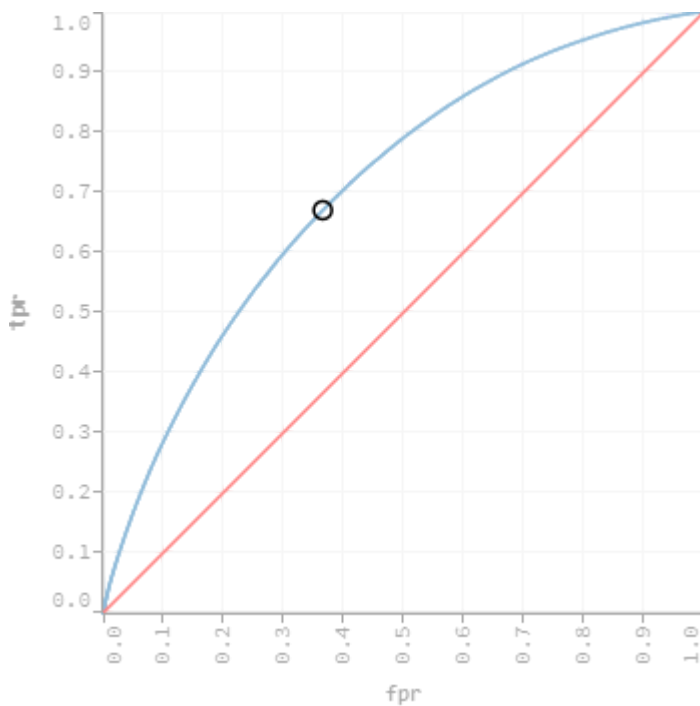
Selected mark(s):

threshold	0.1887
f1	0.4586
f2	0.5657
f0point5	0.3855
accuracy	0.6425
precision	0.3485
recall	0.6701
specificity	0.6345
absolute_mcc	0.2570
min_per_class_accuracy	0.6345
mean_per_class_accuracy	0.6523
tns	878857
fns	133356
fps	506315
tps	270870
tnr	0.6345
fnr	0.3299
fpr	0.3655
tpr	0.6701
idx	257

	Actual/Pred
	0
CM	
	1
	Total

▼ VARIABLE IMPORTANCES

scaled_importance: *1.0*
variable: *precinct*
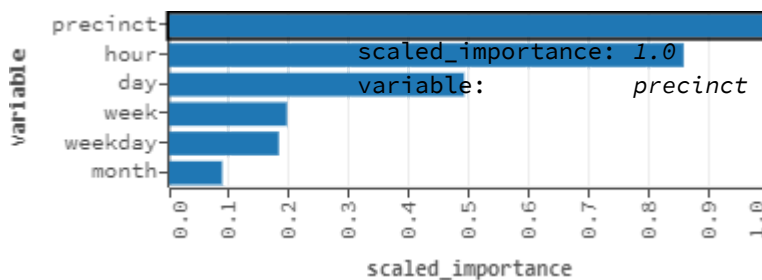
▼ TRAINING METRICS - CONFUSION MATRIX ROW LABELS: ACTUAL CLASS; COLUMN LABELS: PREDICTED CLASS

0	916091	469081	0.3386	469,081 / 1,385,172	0.90
1	101244	1284020	0.0731	101,244 / 1,385,264	0.73
Total	1017335	1753101	0.2059	570,325 / 2,770,436	
Recall	0.66	0.93			

▼ CROSS VALIDATION METRICS - CONFUSION MATRIX ROW LABELS: ACTUAL CLASS; COLUMN LABELS: PREDICTED CLASS

0	878857	506315	0.3655	506,315 / 1,385,172	0.87
1	133356	270870	0.3299	133,356 / 404,226	0.35
Total	1012213	777185	0.3575	639,671 / 1,789,398	
Recall	0.63	0.67			

▾ TRAINING METRICS - GAINS/LIFT TABLE
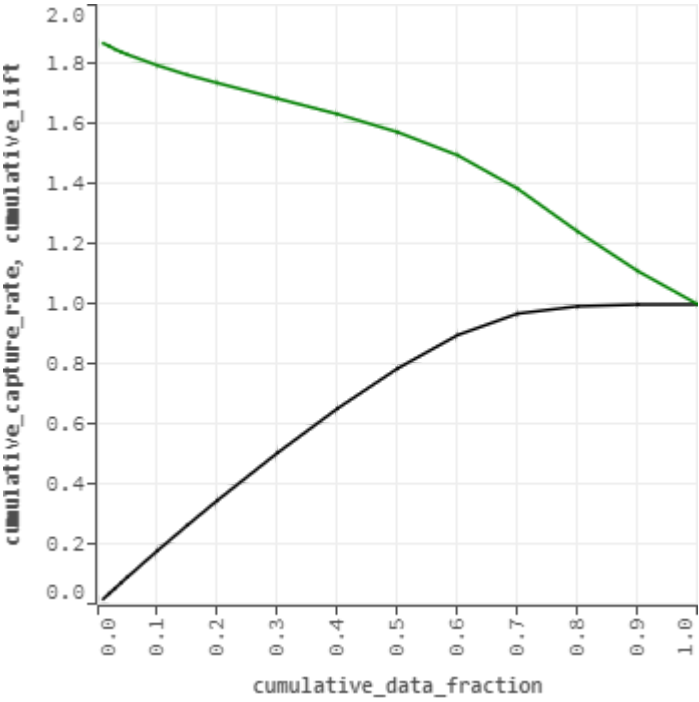
▾ CROSS VALIDATION METRICS - GAINS/LIFT TABLE
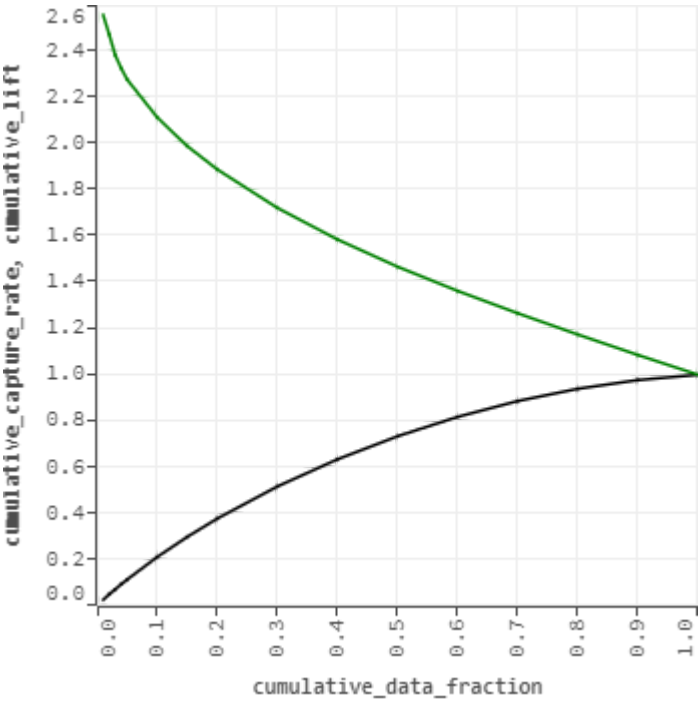
▾ OUTPUT

cross_validation_models	·
cross_validation_holdout_predictions_frame_id	cv_holdout_prediction_DRF_1_AutoML_20200209_085
cross_validation_fold_assignment_frame_id	·
model_category	Binomial
validation_metrics	·
status	·
start_time	1581256757911
end_time	1581256843514
run_time	85603
init_f	0

▸ COLUMN_TYPES

▸ CROSS_VALIDATION_PREDICTIONS

▾ OUTPUT - MODEL SUMMARY

number_of_trees	50
number_of_internal_trees	50
model_size_in_bytes	106675561
min_depth	20
max_depth	20
mean_depth	20.0
min_leaves	113510
max_leaves	158053
mean_leaves	142819.4500

▸ OUTPUT - SCORING HISTORY

▾ OUTPUT - TRAINING_METRICS

model	DRF_1_AutoML_20200209_085106
model_checksum	30940386791445796
frame	automl_training_py_109_sid_96ee
frame_checksum	-4318707167276932420
description	Metrics reported on Out-Of-Bag training samples
model_category	Binomial
scoring_time	1581256843513
predictions	·
MSE	0.212163
RMSE	0.460611
nobs	2770436
custom_metric_name	·
custom_metric_value	0
r2	0.151349
logloss	0.589727
AUC	0.869877
pr_auc	0.835240
Gini	0.739755
mean_per_class_error	0.205866

▼ D O M A I N

domain

0

1

▸ OUTPUT - TRAINING_METRICS - METRICS FOR THRESHOLDS (BINOMIAL METRICS AS A FUNCTION OF CLASSIFICATION THRESHOLDS)

▼ OUTPUT - TRAINING_METRICS - MAXIMUM METRICS (MAXIMUM METRICS AT THEIR RESPECTIVE THRESHOLDS)

metric	threshold	value	idx
max f1	0.2092	0.8183	281
max f2	0.1465	0.8983	312
max f0point5	0.2936	0.7870	234
max accuracy	0.2355	0.7983	266
max precision	0.9992	0.9701	0
max recall	0.0109	1.0	393
max specificity	0.9992	1.0	0
max absolute_mcc	0.2156	0.6103	277
max min_per_class_accuracy	0.2899	0.7852	236
max mean_per_class_accuracy	0.2355	0.7983	266
max tns	0.9992	1385168.0	0
max fns	0.9992	1385134.0	0
max fps	0.0006	1385172.0	399
max tps	0.0109	1385264.0	393
max tnr	0.9992	1.0	0
max fnr	0.9992	0.9999	0
max fpr	0.0006	1.0	399
max tpr	0.0109	1.0	393

▾ OUTPUT - TRAINING_METRICS - GAINS/LIFT TABLE (AVG RESPONSE RATE: 50.00 %, AVG SCORE: 28.88 %)

group	cumulative_data_fraction	lower_threshold	lift	cumulative_lift	response_rate	scor
1	0.0100	0.7423	1.8693	1.8693	0.9347	0.804
2	0.0200	0.6887	1.8506	1.8600	0.9254	0.713
3	0.0300	0.6538	1.8279	1.8493	0.9140	0.670
4	0.0400	0.6274	1.8133	1.8403	0.9067	0.640
5	0.0500	0.6058	1.7986	1.8319	0.8993	0.616
6	0.1000	0.5349	1.7601	1.7960	0.8801	0.566
7	0.1500	0.4899	1.7001	1.7641	0.8501	0.511
8	0.2000	0.4548	1.6567	1.7372	0.8284	0.471
9	0.3000	0.3953	1.5817	1.6854	0.7909	0.424
10	0.4000	0.3421	1.4783	1.6336	0.7392	0.368
11	0.5000	0.2882	1.3331	1.5735	0.6666	0.315
12	0.6000	0.2294	1.1117	1.4965	0.5559	0.259
13	0.7000	0.1605	0.7159	1.3850	0.3580	0.195
14	0.8000	0.0883	0.2437	1.2423	0.1218	0.123
15	0.9000	0.0385	0.0581	1.1108	0.0291	0.061
16	1.0	0	0.0031	1.0	0.0016	0.019

▾ OUTPUT - CROSS_VALIDATION_METRICS

model	DRF_1_AutoML_20200209_085106
model_checksum	30940386791445796
frame	automl_training_py_109_sid_96ee
frame_checksum	-4318707167276932420
description	5-fold cross-validation on training data (Metrics computed for combined holdout predictions)
model_category	Binomial
scoring_time	1581256843667
predictions	·
MSE	0.160310
RMSE	0.400387
nobs	1789398
custom_metric_name	·
custom_metric_value	0
r2	0.083261
logloss	0.500348
AUC	0.708496
pr_auc	0.398261
Gini	0.416992
mean_per_class_error	0.347715

▸ OUTPUT - CROSS_VALIDATION_METRICS - METRICS FOR THRESHOLDS (BINOMIAL METRICS AS A FUNCTION OF CLASSIFICATION THRESHOLDS)

▾ OUTPUT - CROSS_VALIDATION_METRICS - MAXIMUM METRICS (MAXIMUM METRICS AT THEIR RESPECTIVE THRESHOLDS)

metric	*threshold*	*value*	*idx*
max f1	0.1887	0.4586	257
max f2	0.0628	0.6252	345
max f0point5	0.3173	0.4163	177
max accuracy	0.5602	0.7765	61
max precision	0.7532	0.6135	16
max recall	0.0006	1.0	399
max specificity	0.9573	1.0	0
max absolute_mcc	0.2205	0.2584	236
max min_per_class_accuracy	0.1972	0.6507	251
max mean_per_class_accuracy	0.1799	0.6523	262
max tns	0.9573	1385164.0	0
max fns	0.9573	404218.0	0
max fps	0.0006	1385172.0	399
max tps	0.0006	404226.0	399
max tnr	0.9573	1.0	0
max fnr	0.9573	1.0	0
max fpr	0.0006	1.0	399
max tpr	0.0006	1.0	399

▾ OUTPUT - CROSS_VALIDATION_METRICS - GAINS/LIFT TABLE (AVG RESPONSE RATE: 22.59 %, AVG SCORE: 19.30 %)

group	cumulative_data_fraction	lower_threshold	lift	cumulative_lift	response_rate	scor
1	0.0100	0.6251	2.5565	2.5565	0.5775	0.693
2	0.0200	0.5689	2.3900	2.4732	0.5399	0.594
3	0.0300	0.5347	2.2047	2.3837	0.4980	0.550
4	0.0400	0.5101	2.1545	2.3264	0.4867	0.521
5	0.0500	0.4907	2.0810	2.2773	0.4701	0.500
6	0.1000	0.4220	1.9500	2.1137	0.4405	0.453
7	0.1500	0.3733	1.7408	1.9894	0.3932	0.396
8	0.2000	0.3331	1.5854	1.8884	0.3582	0.352
9	0.3000	0.2647	1.3892	1.7220	0.3138	0.298
10	0.4000	0.2060	1.1754	1.5854	0.2655	0.234
11	0.5000	0.1547	0.9964	1.4676	0.2251	0.179
12	0.6000	0.1098	0.8388	1.3628	0.1895	0.131
13	0.7000	0.0735	0.6862	1.2661	0.1550	0.090
14	0.8000	0.0459	0.5301	1.1741	0.1197	0.059
15	0.9000	0.0246	0.3773	1.0856	0.0852	0.034
16	1.0	0	0.2298	1.0	0.0519	0.013

▾ OUTPUT - CROSS-VALIDATION METRICS SUMMARY

	mean	sd	cv_1_valid	cv_2_valid	cv_3_valid	cv_4_vali
accuracy	0.6450309	0.008351063	0.6350928	0.6478401	0.6562535	0.6473417
auc	0.708496	7.211098E-4	0.70749664	0.7091181	0.7083085	0.7082785
aucpr	0.3982794	0.0015449214	0.39632297	0.39963815	0.3981351	0.397313
err	0.3549691	0.008351063	0.36490723	0.35215995	0.3437465	0.3526583
err_count	127036.2	2988.6335	130593.0	126031.0	123020.0	126209.0
f0point5	0.38664234	0.0025114142	0.3834557	0.3875366	0.3900465	0.3871033
f1	0.45871112	8.9054607E-4	0.45932955	0.45842734	0.45750722	0.4584815
f2	0.563913	0.007851634	0.57263637	0.5610602	0.55318344	0.5621337
lift_top_group	2.5573325	0.036490433	2.5164762	2.5974662	2.5797815	2.5209126
logloss	0.5003478	7.311947E-4	0.50161314	0.49984953	0.49984378	0.5001988
max_per_class_error	0.3645265	0.011154508	0.379594	0.35556212	0.3571995	0.3568962
mcc	0.25756449	3.742269E-4	0.25717527	0.25720036	0.25761172	0.2578067
mean_per_class_accuracy	0.65239465	6.7525066E-4	0.6528743	0.6519647	0.6514854	0.6524983
mean_per_class_error	0.34760532	6.7525066E-4	0.3471257	0.34803528	0.34851462	0.3475016
mse	0.16030963	2.3564126E-4	0.16071208	0.16019252	0.16013663	0.1601821
pr_auc	0.3982794	0.0015449214	0.39632297	0.39963815	0.3981351	0.397313
precision	0.34999943	0.0037503298	0.3454174	0.35131824	0.35513598	0.3507039
r2	0.08326065	0.001046671	0.08173737	0.08422545	0.083082825	0.0829853
recall	0.66584194	0.01701947	0.6853426	0.6594916	0.6428005	0.6618929
rmse	0.40038675	2.9415076E-4	0.40088913	0.40024057	0.40017074	0.4002275
specificity	0.6389474	0.015760688	0.62040603	0.64443785	0.66017026	0.6431038

▸ OUTPUT - VARIABLE IMPORTANCES

▸ PREVIEW POJO