

diabetes

?

Assistance

<i>Routine</i>	<i>Description</i>
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: Enter a file or directory path and press the Enter key 

Selected Files: (No files selected)

Actions: Import

?

Assistance

<i>Routine</i>	<i>Description</i>
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

Routine	Description
runAutoML	Automatically train and tune many models
buildModel	Build a model
importModel	Import a saved model
predict	Make a prediction

Import Files

Search: E:\REVA-MBA-BUSINESS-ANALYTICS\TRIMESTER1_Module01_Data



Search Results: (All files added)

Selected Files: 1 file selected: Clear All

✖ E:\REVA-MBA-BUSINESS-ANALYTICS\TRIMESTER1_Module01_DataManagement_and_Visual Analytics-JBS\day2_session\BA06 Module 1 Introduction to Data Management Day 2 Dataset\diabetes.csv

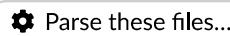
Actions:



1 / 1 files imported.

Files E:\REVA-MBA-BUSINESS-ANALYTICS\TRIMESTER1_Module01_DataManagement_and_Visual Analytics-JBS\day2_session\BA06 Module 1 Introduction to Data Management Day 2 Dataset\diabetes.csv

Actions



Setup Parse

PARSE CONFIGURATION

Sources nfs:E:\REVA-MBA-BUSINESS-ANALYTICS\TRIMESTER1_Module01_DataManagement_and_Visual Analytics-JBS\day2_session\BA06 Module 1 Introduction to Data Management Day 2 Dataset\diabetes.csv

ID diabetes.hex

Parser CSV ▾

Separator ,: '044' ▾

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	Pregnancies	Numeric ▾	6	1	8	1	0	5	3	10	2
2	Glucose	Numeric ▾	148	85	183	89	137	116	78	115	197
3	BloodPressure	Numeric ▾	72	66	64	66	40	74	50	0	70
4	SkinThickness	Numeric ▾	35	29	0	23	35	0	32	0	45
5	Insulin	Numeric ▾	0	0	0	94	168	0	88	0	543
6	BMI	Numeric ▾	33.6	26.6	23.3	28.1	43.1	25.6	31	35.3	30.5
7	DiabetesPed	Numeric ▾	0.627	0.351	0.672	0.167	2.288	0.201	0.248	0.134	0.158
8	Age	Numeric ▾	50	31	32	21	33	30	26	29	53
9	Outcome	Numeric ▾	1	0	1	0	1	0	1	0	1

◀ Previous page

→ Next page

Parse

Job

Run Time 00:00:00.542

Remaining Time 00:00:00.0

Type Frame

Key diabetes.hex

Description Parse

Status DONE

Progress 100%

Done.

Actions

View

diabetes.hex

DATA

◀ Previous 20 Columns

→ Next 20 Columns

78	5.0	95.0	72.0	33.0	0	37.7000
79	0	131.0	0	0	0	43.2000
80	2.0	112.0	66.0	22.0	0	25.0
81	3.0	113.0	44.0	13.0	0	22.4000

Row	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction
83	7.0	83.0	78.0	26.0	71.0	29.3000	
84	0	101.0	65.0	28.0	0	24.6000	
85	5.0	137.0	108.0	0	0	48.8000	
86	2.0	110.0	74.0	29.0	125.0	32.4000	
87	13.0	106.0	72.0	54.0	0	36.6000	
88	2.0	100.0	68.0	25.0	71.0	38.5000	
89	15.0	136.0	70.0	32.0	110.0	37.1000	
90	1.0	107.0	68.0	19.0	0	26.5000	
91	1.0	80.0	55.0	0	0	19.1000	
92	4.0	123.0	80.0	15.0	176.0	32.0	
93	7.0	81.0	78.0	40.0	48.0	46.7000	
94	4.0	134.0	72.0	0	0	23.8000	
95	2.0	142.0	82.0	18.0	64.0	24.7000	
96	6.0	144.0	72.0	27.0	228.0	33.9000	
97	2.0	92.0	62.0	28.0	0	31.6000	
98	1.0	71.0	48.0	18.0	76.0	20.4000	
99	6.0	93.0	50.0	30.0	64.0	28.7000	
100	1.0	122.0	90.0	51.0	220.0	49.7000	

[← Previous 20 Columns](#)[→ Next 20 Columns](#)

diabetes.hex

Actions:

[View Data](#)[Split](#)[Build Model](#)[Run AutoML](#)[Predict](#)[Delete](#)[Download](#)[Export](#)

Rows	Columns	Compressed Size
768	9	10KB

▼ COLUMN SUMMARIES

label	type	Missing	Zeros	+Inf	-Inf	min	max	mean	sigma
Pregnancies	int	0	111	0	0	0	17.0	3.8451	3.3696
Glucose	int	0	5	0	0	0	199.0	120.8945	31.9726
BloodPressure	int	0	35	0	0	0	122.0	69.1055	19.3558
SkinThickness	int	0	227	0	0	0	99.0	20.5365	15.9522
Insulin	int	0	374	0	0	0	846.0	79.7995	115.2440
BMI	real	0	11	0	0	0	67.1000	31.9926	7.8842
DiabetesPedigreeFunction	real	0	0	0	0.0780	2.4200	0.4719	0.3313	

label	type	Missing	Zeros	+Inf	-Inf	min	max	mean	sigma
Age	int	0	0	0	0	21.0	81.0	33.2409	11.7602
Outcome	enum	0	500	0	0	0	1.0	0.3490	0.4770

[◀ Previous 20 Columns](#)[→ Next 20 Columns](#)[▶ CHUNK COMPRESSION SUMMARY](#)[▶ FRAME DISTRIBUTION SUMMARY](#)

✖ Split Frame

Frame: **diabetes.hex ▾**

Splits: Ratio

Key

0.75

frame_0.750

✖

0.250

frame_0.250

Add a new split

Seed: 32003

[✖ Create](#)

田 Split Frames

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

田 frame_0.750

Actions: [View Data](#) [Split](#) [Build Model](#) [Run AutoML](#) [Predict](#) [Delete](#)
[Download](#) [Export](#)

Rows	Columns	Compressed Size
567	9	8KB

[▼ COLUMN SUMMARIES](#)

label	type	Missing	Zeros	+Inf	-Inf	min	max	mean	sigma
-------	------	---------	-------	------	------	-----	-----	------	-------

label	type	Missing	Zeros	+Inf	-Inf	min	max	mean	sigma
Pregnancies	int	0	87	0	0	0	17.0	3.8889	3.4931
Glucose	int	0	4	0	0	0	199.0	120.7478	32.2277
BloodPressure	int	0	24	0	0	0	122.0	69.2840	18.8872
SkinThickness	int	0	166	0	0	0	99.0	20.6208	16.0348
Insulin	int	0	274	0	0	0	846.0	80.0088	114.7451
BMI	real	0	5	0	0	0	67.1000	32.3910	7.5497
DiabetesPedigreeFunction	real	0	0	0	0	0.0780	2.4200	0.4743	0.3329
Age	int	0	0	0	0	21.0	72.0	33.6596	11.8430
Outcome	enum	0	361	0	0	0	1.0	0.3633	0.4814

◀ ▶

[← Previous 20 Columns](#) [→ Next 20 Columns](#)

► CHUNK COMPRESSION SUMMARY

► FRAME DISTRIBUTION SUMMARY

frame_0.250

Actions: [View Data](#) [Split](#) [Build Model](#) [Run AutoML](#) [Predict](#) [Delete](#)
[Download](#) [Export](#)

Rows	Columns	Compressed Size
201	9	4KB

▼ COLUMN SUMMARIES

label	type	Missing	Zeros	+Inf	-Inf	min	max	mean	sigma
Pregnancies	int	0	24	0	0	0	11.0	3.7214	2.9987
Glucose	int	0	1	0	0	0	197.0	121.3085	31.3173
BloodPressure	int	0	11	0	0	0	108.0	68.6020	20.6613
SkinThickness	int	0	61	0	0	0	52.0	20.2985	15.7541
Insulin	int	0	100	0	0	0	680.0	79.2090	116.9266
BMI	real	0	6	0	0	0	57.3000	30.8687	8.6810
DiabetesPedigreeFunction	real	0	0	0	0	0.0840	2.2880	0.4649	0.3275

label	type	Missing	Zeros	+Inf	-Inf	min	max	mean	sigma
Age	int	0	0	0	0	21.0	81.0	32.0597	11.4702
Outcome	enum	0	139	0	0	0	1.0	0.3085	0.4630


[◀ Previous 20 Columns](#)
[→ Next 20 Columns](#)
[▶ CHUNK COMPRESSION SUMMARY](#)
[▶ FRAME DISTRIBUTION SUMMARY](#)

?

 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
.Grid 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

cube Build a Model

Select an algorithm: **Distributed Random Forest** ▾

[PARAMETERS](#)

G

<i>model_id</i>	RANDOM_FOREST_DIABETES	Destination id for this model; auto-generated if not specified.
<i>training_frame</i>	frame_0.750 ▾	Id of the training data frame.
<i>validation_frame</i>	frame_0.250 ▾	Id of the validation data frame.

<i>nfold</i>	0	Number of folds for K-fold cross-validation (0 to disable or >= 2).																				
<i>response_column</i>	Outcome	Response variable column.																				
<i>ignored_columns</i>	Search...	Names of columns to ignore for training.																				
<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="padding: 2px;"><input type="checkbox"/> Pregnancies</td> <td style="padding: 2px;">INT</td> </tr> <tr> <td style="padding: 2px;"><input type="checkbox"/> Glucose</td> <td style="padding: 2px;">INT</td> </tr> <tr> <td style="padding: 2px;"><input type="checkbox"/> BloodPressure</td> <td style="padding: 2px;">INT</td> </tr> <tr> <td style="padding: 2px;"><input type="checkbox"/> SkinThickness</td> <td style="padding: 2px;">INT</td> </tr> <tr> <td style="padding: 2px;"><input type="checkbox"/> Insulin</td> <td style="padding: 2px;">INT</td> </tr> <tr> <td style="padding: 2px;"><input type="checkbox"/> BMI</td> <td style="padding: 2px;">REAL</td> </tr> <tr> <td style="padding: 2px;"><input type="checkbox"/></td> <td style="padding: 2px;">REAL</td> </tr> <tr> <td colspan="2" style="text-align: center; padding: 2px;">DiabetesPedigreeFunction</td> </tr> <tr> <td style="padding: 2px;"><input type="checkbox"/> Age</td> <td style="padding: 2px;">INT</td> </tr> <tr> <td style="padding: 2px;"><input type="checkbox"/> Outcome</td> <td style="padding: 2px;">ENUM(2)</td> </tr> </table>		<input type="checkbox"/> Pregnancies	INT	<input type="checkbox"/> Glucose	INT	<input type="checkbox"/> BloodPressure	INT	<input type="checkbox"/> SkinThickness	INT	<input type="checkbox"/> Insulin	INT	<input type="checkbox"/> BMI	REAL	<input type="checkbox"/>	REAL	DiabetesPedigreeFunction		<input type="checkbox"/> Age	INT	<input type="checkbox"/> Outcome	ENUM(2)	
<input type="checkbox"/> Pregnancies	INT																					
<input type="checkbox"/> Glucose	INT																					
<input type="checkbox"/> BloodPressure	INT																					
<input type="checkbox"/> SkinThickness	INT																					
<input type="checkbox"/> Insulin	INT																					
<input type="checkbox"/> BMI	REAL																					
<input type="checkbox"/>	REAL																					
DiabetesPedigreeFunction																						
<input type="checkbox"/> Age	INT																					
<input type="checkbox"/> Outcome	ENUM(2)																					
<input checked="" type="radio"/> All <input type="radio"/> None																						
Only show columns with more than <input type="text" value="0"/> % missing values.																						
<i>ignore_const_cols</i>	<input checked="" type="checkbox"/>	Ignore constant columns.																				
<i>ntrees</i>	50	Number of trees.																				
<i>max_depth</i>	20	Maximum tree depth.																				
<i>min_rows</i>	1	Fewest allowed (weighted) observations in a leaf.																				
<i>nbins</i>	20	For numerical columns (real/int), build a histogram of (at least) this many bins, then split at the best point																				
<i>seed</i>	-1	Seed for pseudo random number generator (if applicable)																				
<i>mtries</i>	-1	Number of variables randomly sampled as candidates at each split. If set to -1, defaults to \sqrt{p} for classification and $p/3$ for regression (where p is the # of predictors)																				
<i>sample_rate</i>	0.632	Row sample rate per tree (from 0.0 to 1.0)																				

<code>score_each_iteration</code>	<input type="checkbox"/>	Whether to score during each iteration of model training.
<code>score_tree_interval</code>	0	Score the model after every so many trees. Disabled if set to 0.
<code>fold_column</code>	(Choose...)	Column with cross-validation fold index assignment per observation.
<code>offset_column</code>	(Choose...)	Offset column. This will be added to the combination of columns before applying the link function.
<code>weights_column</code>	(Choose...)	Column with observation weights. Giving some observation a weight of zero is equivalent to excluding it from the dataset; giving an observation a relative weight of 2 is equivalent to repeating that 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 typically 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.
<code>balance_classes</code>	<input type="checkbox"/>	Balance training data class counts via over/under-sampling (for imbalanced data).
<code>nbins_top_level</code>	1024	For numerical columns (real/int), build a histogram of (at most) this many bins at the root level, then decrease by factor of two per level
<code>nbins_cats</code>	1024	For categorical columns (factors), build a histogram of this many bins, then split at the best point. Higher values can lead to more overfitting.

`r2_stopping` 1.7976931348623157e+308

`r2_stopping` is no longer supported and will be ignored if set - please use `stopping_rounds`, `stopping_metric` and `stopping_tolerance` instead.

Previous version of H2O would stop making trees when the R^2 metric equals or exceeds this

`stopping_rounds` 0

Early stopping based on convergence of `stopping_metric`. Stop if simple moving average of length k of the `stopping_metric` does not improve for $k:=stopping_rounds$ scoring events (0 to disable)

`stopping_metric` AUTO ▾

Metric to use for early stopping (AUTO: logloss for classification, deviance for regression and anomaly_score for Isolation Forest). Note that custom and custom_increasing can only be used in GBM and DRF with the Python client.

`stopping_tolerance` 0.001

Relative tolerance for metric-based stopping criterion (stop if relative improvement is not at least this much)

`max_runtime_secs` 0

Maximum allowed runtime in seconds for model training. Use 0 to disable.

`checkpoint`

Model checkpoint to resume training with.

`col_sample_rate_per_tree` 1

Column sample rate per tree (from 0.0 to 1.0)

`min_split_improvement` 0.00001

Minimum relative improvement in squared error reduction for a split to happen

`histogram_type` AUTO ▾

What type of histogram to use for finding optimal split points

`categorical_encoding` AUTO ▾

Encoding scheme for categorical features

`distribution` AUTO ▾

Distribution function

`custom_metric_func`

Reference to custom evaluation function, format:
'language:keyName=funcName'

`export_checkpoints_dir`

Automatically export generated models to this directory.

EXPERT

G

`build_tree_one_node`

Run on one node only; no network overhead but fewer cpus used. Suitable for small datasets.

`sample_rate_per_class`

A list of row sample rates per class (relative fraction for each class, from 0.0 to 1.0), for each tree

`binomial_double_trees`

For binary classification: Build 2x as many trees (one per class) - can lead to higher accuracy.

`col_sample_rate_change_per_level` 1

Relative change of the column sampling rate for every level (must be > 0.0 and <= 2.0)

`calibrate_model`

Use Platt Scaling to calculate calibrated class probabilities. Calibration can provide more accurate estimates of class probabilities.

`calibration_frame`

Calibration frame for Platt Scaling

`check_constant_response`

Check if response column is constant. If enabled, then an exception is thrown if the response column is a constant value. If disabled, then model will train regardless of the response column being a constant value or not.

 Build Model

☰ Job

Run Time 00:00:04.580

Remaining Time 00:00:00.0

Type Model

Key  RANDOM_FOREST_DIABETES

Description DRF

Status DONE

Progress 100%

Done.

Actions

View

Model

Model ID: RANDOM_FOREST_DIABETES

Algorithm: Distributed Random Forest

Actions:

Refresh

Predict...

Download POJO

Download Model Deployment Package (MOJO)

Export

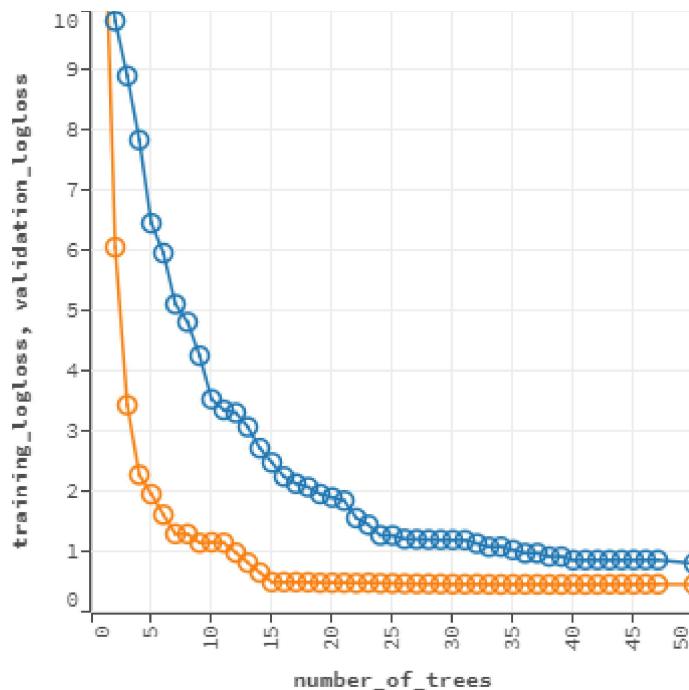
Inspect

Delete

Download Gen Model

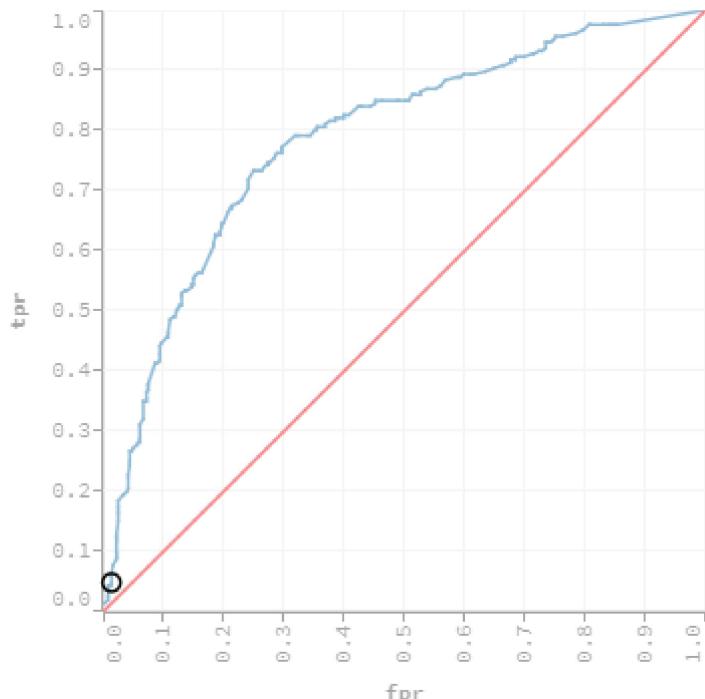
► MODEL PARAMETERS

▼ SCORING HISTORY - LOGLOSS



▼ ROC CURVE - TRAINING METRICS , AUC = 0.785312

H2O- diabetes

**Selected mark(s):**

threshold 0.9000
 f1 0.0905
 f2 0.0596
 f0point5 0.1880
 accuracy 0.6455
 precision 0.6667
 recall 0.0485
 specificity 0.9861
 absolute_mcc 0.1040
 min_per_class_accuracy 0.0485
 mean_per_class_accuracy 0.5173
 tns 356
 fns 196
 fps 5
 tps 10
 tnr 0.9861
 fnr 0.9515
 fpr 0.0139
 tpr 0.0485
 idx 8

Threshold:

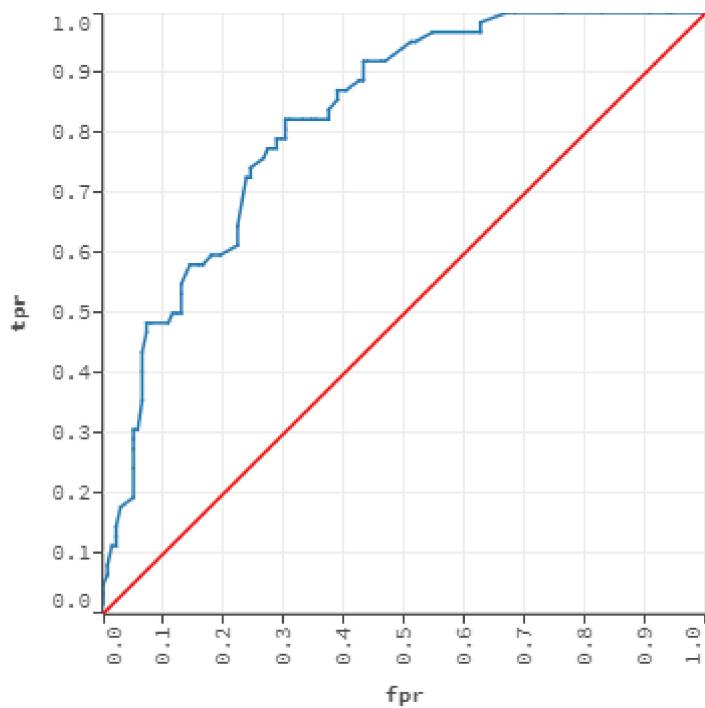
0.375

Criterion:

▼ Choose... ▼

Actual/Predict	
CM	Total
0	356
1	196

▼ ROC CURVE - VALIDATION METRICS , AUC = 0.824959

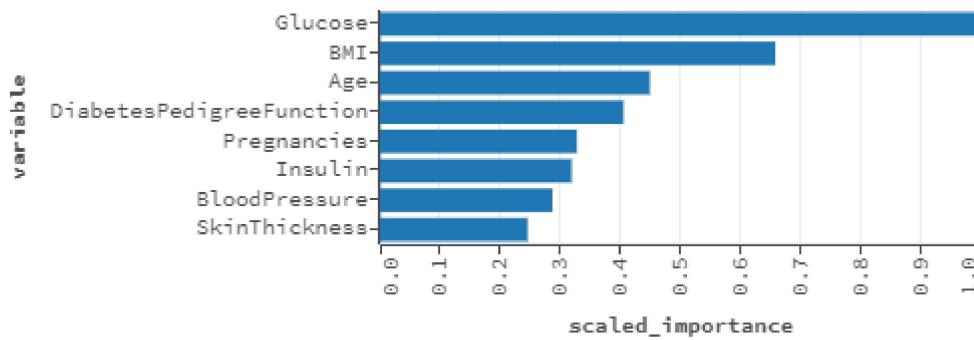
**Threshold:**

Choose...

Criterion:

▼ Choose... ▼

▼ VARIABLE IMPORTANCES



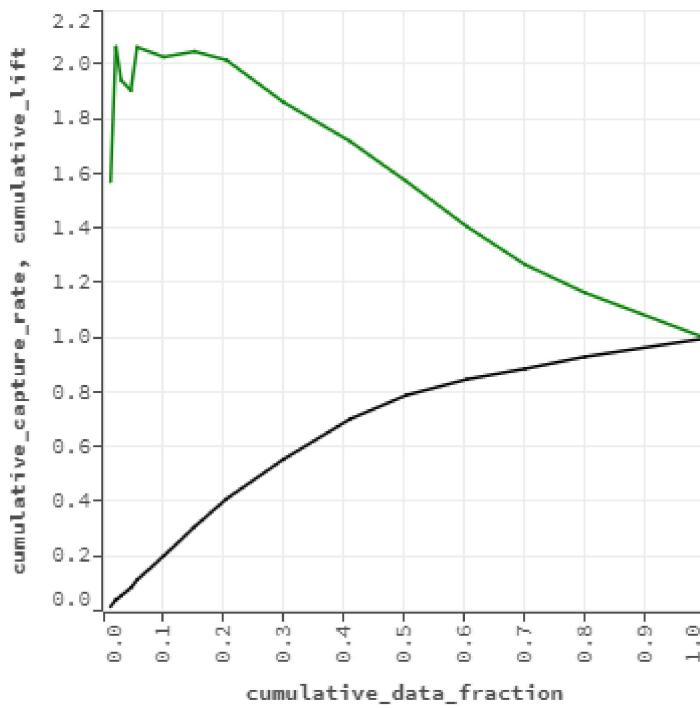
▼ TRAINING METRICS - CONFUSION MATRIX ROW LABELS: ACTUAL CLASS; COLUMN LABELS: PREDICTED CLASS

	0	1	Error	Rate	Precision	▲
0	271	90	0.2493	90 / 361	0.83	▼
1	55	151	0.2670	55 / 206	0.63	▼
Total	326	241	0.2557	145 / 567		▼
Recall	0.75	0.73				▼

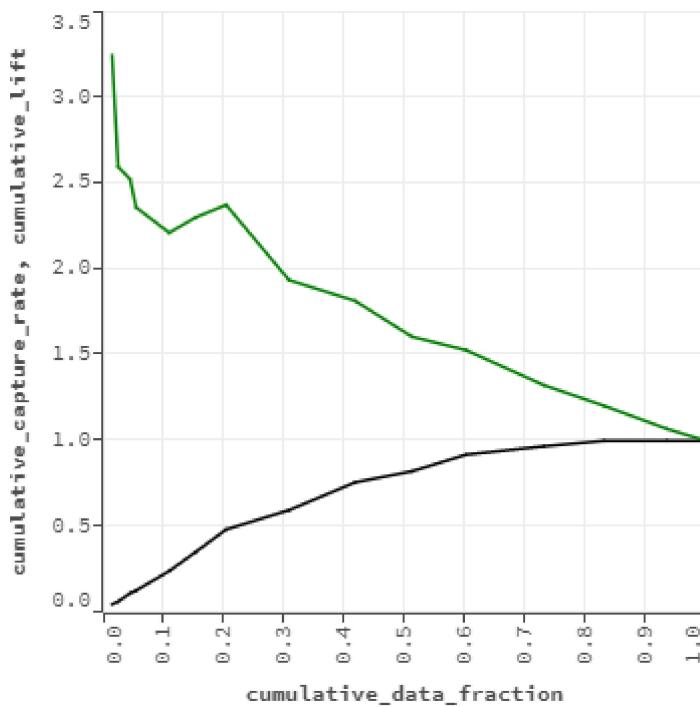
▼ VALIDATION METRICS - CONFUSION MATRIX ROW LABELS: ACTUAL CLASS; COLUMN LABELS: PREDICTED CLASS

	0	1	Error	Rate	Precision	▲
0	97	42	0.3022	42 / 139	0.90	▼
1	11	51	0.1774	11 / 62	0.55	▼
Total	108	93	0.2637	53 / 201		▼
Recall	0.70	0.82				▼

▼ TRAINING METRICS - GAINS/LIFT TABLE



▼ VALIDATION METRICS - GAINS/LIFT TABLE



► OUTPUT

► COLUMN_TYPES

► OUTPUT - MODEL SUMMARY

► OUTPUT - SCORING HISTORY

► OUTPUT - TRAINING_METRICS

► DOMAIN

▶ 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)

▶ OUTPUT - TRAINING_METRICS - GAINS/LIFT TABLE (AVG RESPONSE RATE: 36.33 %, AVG SCORE: 36.59 %)

▶ OUTPUT - VALIDATION_METRICS

▶ OUTPUT - VALIDATION_METRICS - METRICS FOR THRESHOLDS (BINOMIAL METRICS AS A FUNCTION OF CLASSIFICATION THRESHOLDS)

▶ OUTPUT - VALIDATION_METRICS - MAXIMUM METRICS (MAXIMUM METRICS AT THEIR RESPECTIVE THRESHOLDS)

▶ OUTPUT - VALIDATION_METRICS - GAINS/LIFT TABLE (AVG RESPONSE RATE: 30.85 %, AVG SCORE: 34.63 %)

▶ OUTPUT - VARIABLE IMPORTANCES

▼ PREVIEW POJO

 Preview POJO

⚡ Predict

Name:

Model:

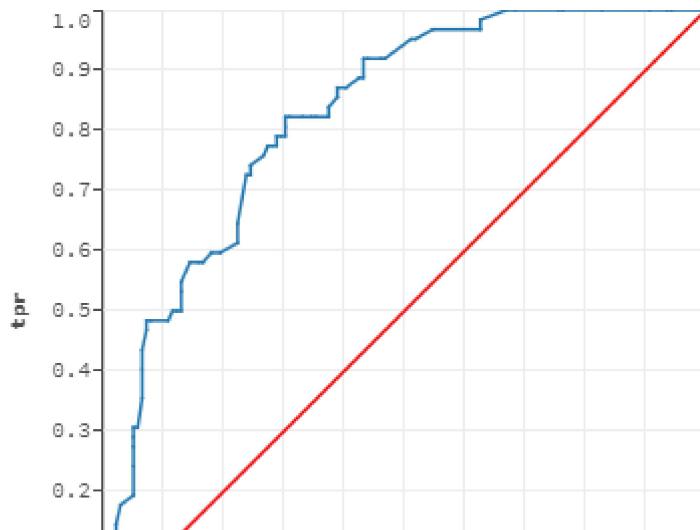
Frame: frame_0.250

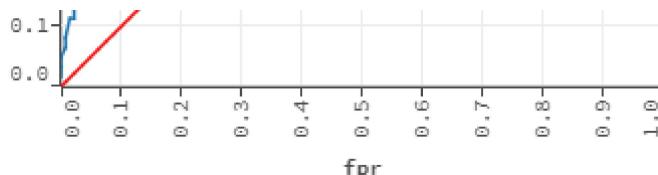
Actions:  Predict

⚡ Prediction

Actions:  Inspect

▼ ROC CURVE





▼ PREDICTION

```

model      RANDOM_FOREST_DIABETES
model_checksum -5823313923050434114
frame      frame_0.250
frame_checksum -2712872429023292051
description .
model_category Binomial
scoring_time 1587194445403
predictions prediction
MSE 0.156574
RMSE 0.395694
nobs 201
custom_metric_name .
custom_metric_value 0
r2 0.265984
logloss 0.467786
AUC 0.824959
pr_auc 0.648736
Gini 0.649919
mean_per_class_error 0.239789

```

Combine predictions with frame

▼ DOMAIN

domain
0
1

▼ PREDICTION - CONFUSION MATRIX ROW LABELS: ACTUAL CLASS; COLUMN LABELS: PREDICTED CLASS

	0	1	Error	Rate	Precision
0	97	42	0.3022	42 / 139	0.90
1	11	51	0.1774	11 / 62	0.55
Total	108	93	0.2637	53 / 201	
Recall	0.70	0.82			

▼ PREDICTION - METRICS FOR THRESHOLDS

threshold	f1	f2	f0point5	accuracy	precision	recall	specificity	absolute_mcc
-----------	----	----	----------	----------	-----------	--------	-------------	--------------

0.9600	0.0317	0.0201	0.0758	0.6965	1.0	0.0161	1.0	0.1059
--------	--------	--------	--------	--------	-----	--------	-----	--------

threshold	f1	f2	f0point5	accuracy	precision	recall	specificity	absolute_mcc
0.9300	0.0625	0.0400	0.1429	0.7015		1.0	0.0323	1.0
0.9000	0.0923	0.0598	0.2027	0.7065		1.0	0.0484	1.0
0.8800	0.1194	0.0791	0.2439	0.7065		0.8000	0.0645	0.9928
0.8700	0.1471	0.0984	0.2907	0.7114		0.8333	0.0806	0.9928
0.8600	0.1748	0.1177	0.3375	0.7162		0.8667	0.1067	0.9928
0.8500	0.2025	0.1364	0.3843	0.7210		0.8889	0.1333	0.9928
0.8400	0.2302	0.1551	0.4311	0.7258		0.9111	0.1600	0.9928
0.8300	0.2579	0.1738	0.4779	0.7306		0.9333	0.1867	0.9928
0.8200	0.2856	0.1925	0.5247	0.7354		0.9556	0.2133	0.9928
0.8100	0.3133	0.2112	0.5715	0.7402		0.9778	0.2500	0.9928
0.8000	0.3410	0.2299	0.6183	0.7450		0.9900	0.2867	0.9928
0.7900	0.3687	0.2486	0.6651	0.7498		0.9922	0.3233	0.9928
0.7800	0.3964	0.2673	0.7119	0.7546		0.9944	0.3600	0.9928
0.7700	0.4241	0.2860	0.7587	0.7594		0.9966	0.3967	0.9928
0.7600	0.4518	0.3047	0.8055	0.7642		0.9988	0.4333	0.9928
0.7500	0.4795	0.3234	0.8523	0.7690		0.9999	0.4700	0.9928
0.7400	0.5072	0.3421	0.8991	0.7738		1.0	0.5067	0.9928
0.7300	0.5349	0.3608	0.9459	0.7786				

▼ PREDICTION - MAXIMUM METRICS

metric	threshold	value	idx
max f1	0.3300	0.6581	46
max f2	0.1400	0.7813	64
max f0point5	0.6300	0.6757	20
max accuracy	0.6300	0.7910	20
max precision	0.9600	1.0	0
max recall	0.1000	1.0	67
max specificity	0.9600	1.0	0
max absolute_mcc	0.3300	0.4821	46
max min_per_class_accuracy	0.4200	0.7419	37
max mean_per_class_accuracy	0.3300	0.7602	46
max tns	0.9600	139.0	0
max fns	0.9600	61.0	0
max fps	0	139.0	75
max tps	0.1000	62.0	67
max tnr	0.9600	1.0	0
max fnr	0.9600	0.9839	0
max fpr	0	1.0	75
max tpr	0.1000	1.0	67

▼ PREDICTION - GAINS/LIFT TABLE

group	cumulative_data_fraction	lower_threshold	lift	cumulative_lift	response_rate
1	0.0149	0.9000	3.2419	3.2419	1.0 0
2	0.0249	0.8800	1.6210	2.5935	0.5000 0
3	0.0448	0.8600	2.4315	2.5215	0.7500 0
4	0.0547	0.8400	1.6210	2.3578	0.5000 0
5	0.1095	0.7800	2.0630	2.2104	0.6364 0

group	cumulative_data_fraction	lower_threshold	lift	cumulative_lift	response_rate	
6	0.1542	0.7000	2.5215	2.3007	0.7778	0
7	0.2040	0.6200	2.5935	2.3721	0.8000	0
8	0.3085	0.4800	1.0806	1.9347	0.3333	0
9	0.4179	0.4100	1.4736	1.8139	0.4545	0
10	0.5124	0.2800	0.6825	1.6052	0.2105	0
11	0.6020	0.1800	1.0806	1.5272	0.3333	0
12	0.7313	0.1200	0.3741	1.3232	0.1154	0
13	0.8308	0.0800	0.3242	1.2036	0.1000	0
14	0.9353	0.0400	0	1.0691	0	0
15	1.0	0	0	1.0	0	0

grid Frames Combined

The specified frames were combined successfully.

[View Frame](#)

CS

grid combined-prediction

Actions: [View Data](#) [Split](#) [Build Model](#) [Run AutoML](#) [Predict](#) [Delete](#)
[Download](#) [Export](#)

Rows	Columns	Compressed Size
201	12	7KB

▼ COLUMN SUMMARIES

label	type	Missing	Zeros	+Inf	-Inf	min	max	mean	sig
predict	enum	0	108	0	0	0	1.0	0.4627	0.45
p0	real	0	0	0	0	0.0400	1.0	0.6537	0.27
p1	real	0	8	0	0	0	0.9600	0.3463	0.27
Pregnancies	int	0	24	0	0	0	11.0	3.7214	2.95
Glucose	int	0	1	0	0	0	197.0	121.3085	31.31
BloodPressure	int	0	11	0	0	0	108.0	68.6020	20.66
SkinThickness	int	0	61	0	0	0	52.0	20.2985	15.75
Insulin	int	0	100	0	0	0	680.0	79.2090	116.92
BMI	real	0	6	0	0	0	57.3000	30.8687	8.65

label	type	Missing	Zeros	+Inf	-Inf	min	max	mean	size
DiabetesPedigreeFunction	real	0	0	0	0	0.0840	2.2880	0.4649	0.32
Age	int	0	0	0	0	21.0	81.0	32.0597	11.47
Outcome	enum	0	139	0	0	0	1.0	0.3085	0.46

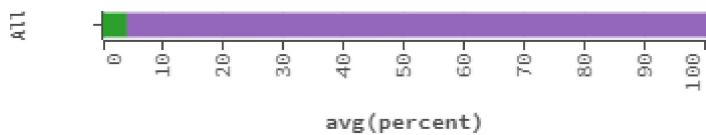
[◀ Previous 20 Columns](#)[▶ Next 20 Columns](#)[▶ CHUNK COMPRESSION SUMMARY](#)[▶ FRAME DISTRIBUTION SUMMARY](#)

Summary: p1

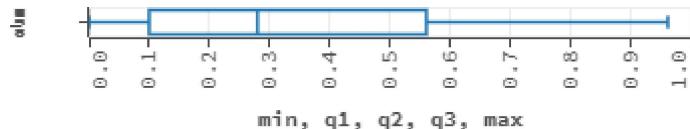
Actions

[◀ Impute](#)[☰ Inspect](#)

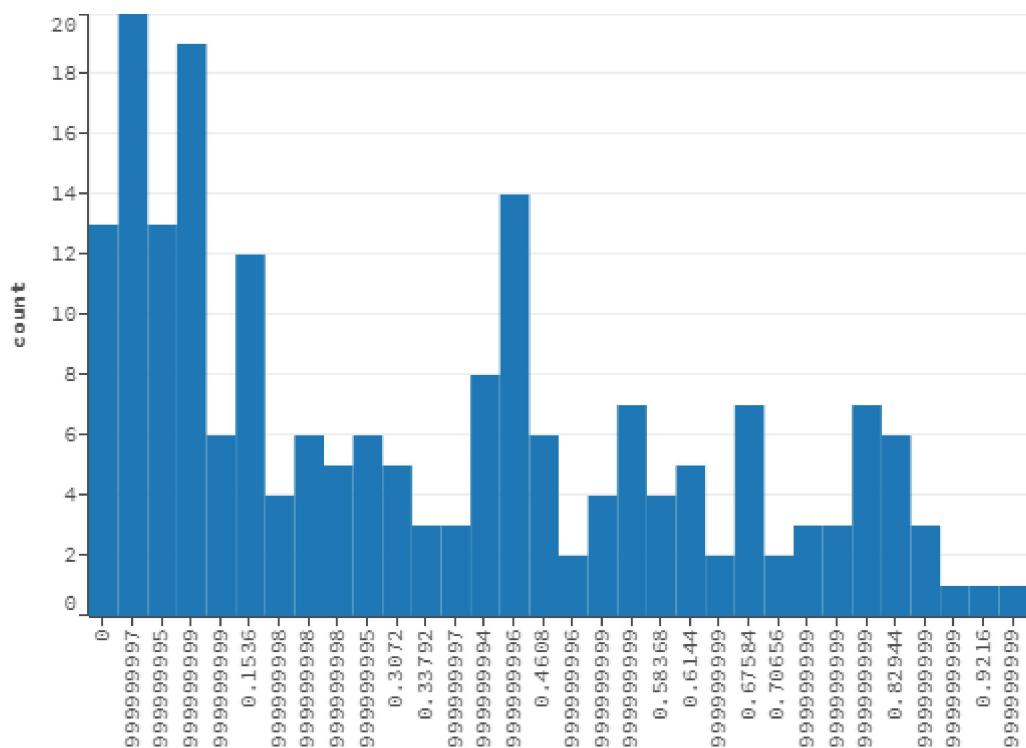
CHARACTERISTICS

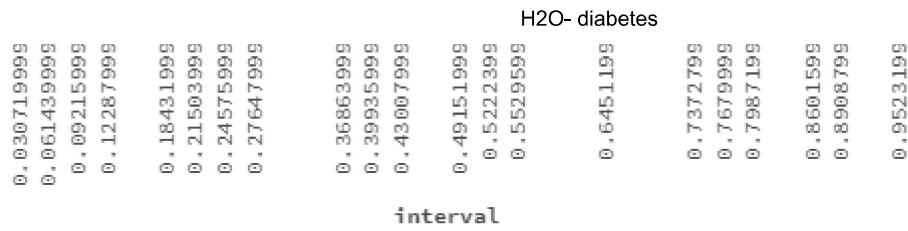


SUMMARY



DISTRIBUTION

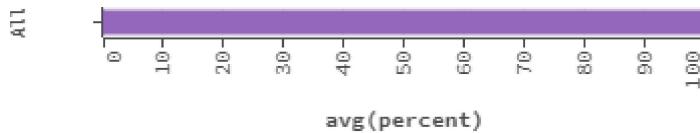




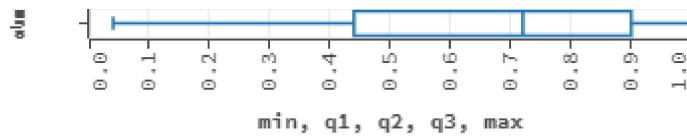
➊ Summary: p0

Actions

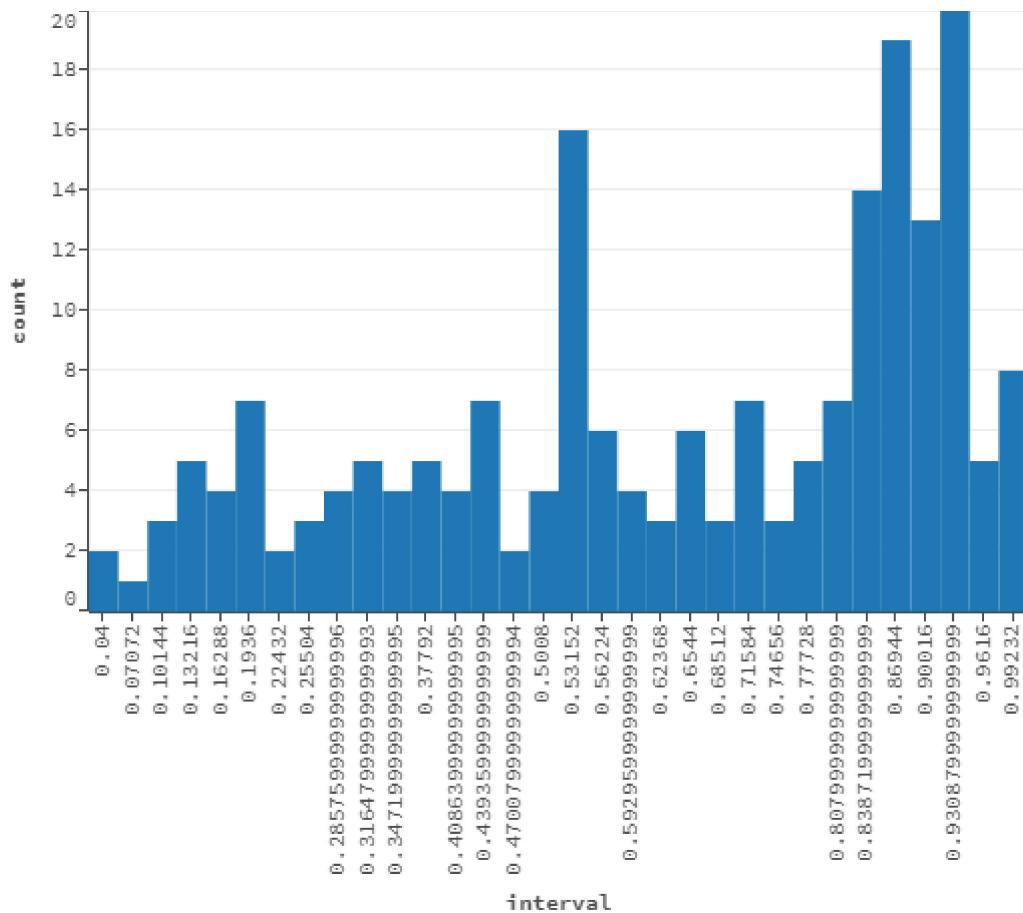
CHARACTERISTICS



SUMMARY



DISTRIBUTION



➋ combined-prediction

▼ DATA

[◀ Previous 20 Columns](#)[▶ Next 20 Columns](#)

Row	predict	p0	p1	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	
1	0	0.8800	0.1200	1.0	85.0	66.0	29.0	0	2
2	1	0.3400	0.6600	0	137.0	40.0	35.0	168.0	4
3	1	0.4300	0.5700	10.0	115.0	0	0	0	3
4	1	0.3000	0.7000	2.0	197.0	70.0	45.0	543.0	3
5	0	0.8600	0.1400	8.0	125.0	96.0	0	0	
6	1	0.5900	0.4100	7.0	100.0	0	0	0	
7	0	0.7400	0.2600	1.0	103.0	30.0	38.0	83.0	4
8	0	0.8400	0.1600	1.0	115.0	70.0	30.0	96.0	3
9	1	0.1700	0.8300	7.0	196.0	90.0	0	0	3
10	0	1.0	0	1.0	97.0	66.0	15.0	140.0	2
11	0	0.9400	0.0600	6.0	92.0	92.0	0	0	1
12	1	0.2000	0.8000	3.0	180.0	64.0	25.0	70.0	
13	1	0.4350	0.5650	1.0	146.0	56.0	0	0	2
14	1	0.5500	0.4500	7.0	103.0	66.0	32.0	0	3
15	0	0.9800	0.0200	1.0	101.0	50.0	15.0	36.0	2
16	1	0.1200	0.8800	7.0	150.0	66.0	42.0	342.0	3
17	1	0.1300	0.8700	7.0	187.0	68.0	39.0	304.0	3
18	1	0.4600	0.5400	0	146.0	82.0	0	0	4
19	1	0.5600	0.4400	0	109.0	88.0	30.0	0	3
20	0	0.8900	0.1100	2.0	100.0	66.0	20.0	90.0	3
21	1	0.4800	0.5200	5.0	139.0	64.0	35.0	140.0	2
22	0	0.7800	0.2200	5.0	95.0	72.0	33.0	0	3
23	0	0.9600	0.0400	2.0	112.0	66.0	22.0	0	

[◀ Previous 20 Columns](#)[▶ Next 20 Columns](#)