

MAADSBML AutoML Report For OTICS ADVANCED ANALYTICS

Generated On: 2024-07-14 19:16:56 (UTC)

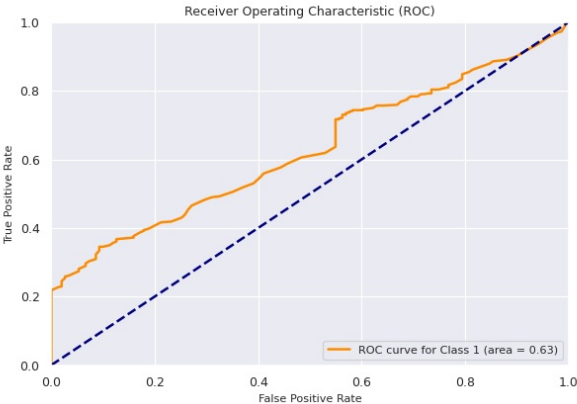
Best Model(s) Report For admin_aesopowerdemandlogistic_csv

MODEL DESCRIPTION

Model Trained On: 2024/07/14
Training Start Time: 1913
Training End Time: 1914
Was Data Normalized: Yes
Was Data Shuffled: Yes
Deep Analysis: No
Total Training Data Set: 961
Training Data Percentage: 75%
Total Test Data Set: 319
Total # of Variables: 4
Adjusted for Seasonality: N
Total Algorithms Run: 450
Removed Outliers: N
ROC AUC: 0.500
Precision: 0.781 (0.308 For Class=0)
Recall: 0.659 (0.450 For Class=0)
F1-Score: 0.715 (0.366 For Class=0)
Best Distribution FOR ACTUAL Y: [RECIPINVGAUSS](#)
Dependent Variable: AESO_POWER_DEMAND_LABEL
Independent Variables: ['Calgary_Weather', 'Edmonton_Weather', 'FtMac_Weather']

Receiver Operating Characteristic Curve (ROC)

Using VotingClassifier algorithm for allseason



IMPORTANT FILE PATHS FOR RAW AND OUTPUT DATA

NOTE: These are DOCKER CONTAINER Paths. You can view these files inside the container by using the command: `docker exec -it {container id} bash` If you have re-run the container, these files will be GONE but they exist on your HOST machine. The HOST MACHINE location is based on the volumes you mapped when you ran the Docker container. The Docker RUN Volume Mappings are :: (For example here is the docker run command (use multiple -v for multiple mappings):

DOCKER RUN COMMAND: `docker run -d -p 5595:5595 -p 5495:5495 -p 10000:10000 -v {HOST MACHINE FOLDER}:{CONTAINER FOLDER}:z --env TRAININGPORT=5595 --env PREDICTIONPORT=5495 --env ABORTPORT=10000 --env COMPANYNAME=MYCOMPANY --env MAXRUNTIME=20 --env MAINHOST=127.0.0.1 maadsdocker/maads-batch-automi-otics`

Docker Volume Mappings:
1. {HOST MACHINE FOLDER}/csvuploads:/maads/agentfilesdocker/dist/maadsweb/csvuploads:z
2. {HOST MACHINE FOLDER}/pdfreports:/maads/agentfilesdocker/dist/maadsweb/pdfreports:z
3. {HOST MACHINE FOLDER}/autofeatures:/maads/agentfilesdocker/dist/maadsweb/autofeatures:z
4. {HOST MACHINE FOLDER}/outliers:/maads/agentfilesdocker/dist/maadsweb/outliers:z
5. {HOST MACHINE FOLDER}/sqlloads:/maads/agentfilesdocker/dist/maadsweb/sqlloads:z
6. {HOST MACHINE FOLDER}/networktemp:/maads/agentfilesdocker/dist/maadsweb/networktemp:z
7. {HOST MACHINE FOLDER}/networks:/maads/agentfilesdocker/networks:z
8. {HOST MACHINE FOLDER}/exception:/maads/agentfilesdocker/dist/maadsweb/exception:z
9. {HOST MACHINE FOLDER}/staging:/maads/agentfilesdocker/dist/staging:z

Path for Training Dataset File: [/maads/agentfilesdocker/dist/maadsweb/csvuploads/aesopowerdemandlogistic_csv](#)
Path for PDF Report (i.e. this file): [/maads/agentfilesdocker/dist/maadsweb/pdfreports/admin_aesopowerdemandlogistic_csv_no_seasons.pdf](#)
Path for AutoFeature File: [/maads/agentfilesdocker/dist/maadsweb/autofeatures/admin_aesopowerdemandlogistic_csv_csv](#)
Path for Outliers File: [/maads/agentfilesdocker/dist/maadsweb/outliers/admin_aesopowerdemandlogistic_csv_csv](#)
Path for Algo JSON File: [/maads/agentfilesdocker/dist/maadsweb/exception/admin_aesopowerdemandlogistic_csv_trained_algo_no_seasons.json](#)
Folder Path for MySQL Scripts: [/maads/agentfilesdocker/dist/maadsweb/sqlloads/](#)
Path for Detailed Prediction File: [/maads/agentfilesdocker/dist/maadsweb/csvuploads/admin_aesopowerdemandlogistic_csv_prediction_details.csv](#)
Path for Algorithm Zip File (i.e pickle files): [/maads/agentfilesdocker/dist/maadsweb/networktemp/admin_aesopowerdemandlogistic_csv.zip](#)
Path for Algorithm Pickle Files:
1. [/maads/agentfilesdocker/networks/Otics Advanced Analytics_ADMIN_AESOPOWERDEMANDLOGISTIC_CSVALLEASON_AG1_4_VotingClassifier_normal_961_ensemble.pkl](#)
2. [/maads/agentfilesdocker/networks/Otics Advanced Analytics_ADMIN_AESOPOWERDEMANDLOGISTIC_CSVALLEASON_AG1_4_VotingClassifier_normal_961_ensemble_scalerx.pkl](#)

DESCRIPTIVE STATISTICS

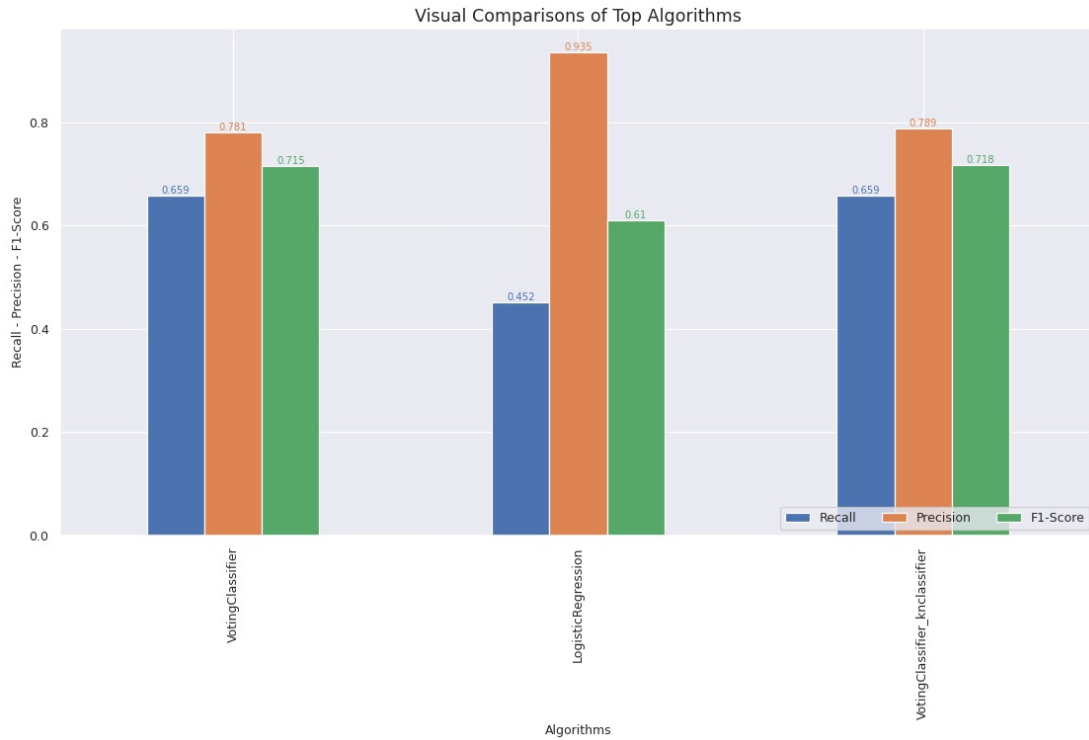
Variables	T-Statistic	Coefficients (alpha,beta)	Count	Mean	STD	MIN	25%	50%	75%	MAX
Calgary_Weather	-3.64	0.716, -0.005	961.0	5.862	-27.75	10.113	-0.4	6.85	14.0	23.85
Edmonton_Weather	-3.505	0.713, -0.004	961.0	5.916	-26.64	11.759	-2.25	7.16	16.1	25.75
FtMac_Weather	-3.116	0.698, -0.003	961.0	2.367	-32.4	13.694	-7.65	4.56	14.8	23.85
AESO_POWER_DEMAND_LABEL	NA	NA	961.0	0.591	0.0	0.492	0.0	1.0	1.0	1.0

BEST ALGORITHM FOUND FOR THIS DATASET

(Note: This trained model will be used to predict [AESO_POWER_DEMAND_LABEL](#))

Algorithm	Description	Model Results	ROC/AUC	Precision	Recall	F1-Score	Forecast Months	Season
VotingClassifier	Voting Classifier: Combination of different classifiers (DecisionTree, RandomForest, K nearest neighbour, GaussNB, Extra tree, ADA boost, etc)	VotingClassifier(estimators=[('dt', DecisionTreeClassifier(class_weight='balanced', max_features='sqrt', max_leaf_nodes=5, min_impurity_decrease=0.027567717177601293, min_samples_leaf=3, min_samples_split=4, min_weight_fraction_leaf=0.45761701969637497, random_state=25, splitter='random')), ('knn', KNeighborsClassifier(algorithm='brute', leaf_size=3, n_neighbors... ('gau', GaussianNB())), ('ab', AdaBoostClassifier(learning_rate=8.905680376428858, n_estimators=12, random_state=70)), ('et', ExtraTreesClassifier(class_weight='balanced', criterion='entropy', max_depth=8, min_impurity_decrease=0.9140453773873057, min_samples_leaf=3, min_samples_split=7, min_weight_fraction_leaf=0.14786636911285872, n_estimators=58, random_state=193))], voting='soft')	0.500	0.781: Class=1 (0.308: Class=0)	0.659: Class=1 (0.450: Class=0)	0.715: Class=1 (0.366: Class=0)	1 - 12	allseason

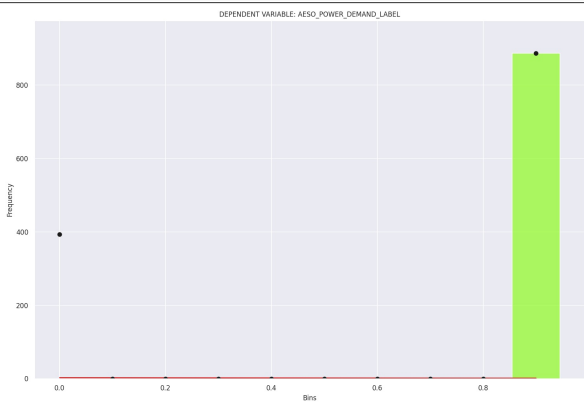
TOP 10 ALGORITHMS FOR ALLSEASON



Num	Algorithm	Model ROC/AUC	Details	Season	Description
1	VotingClassifier	0.5000	Recall: 0.659 (class 1) Precision: 0.781 (class 1) F1 Score: 0.715 (class 1) Recall: 0.45 (class 0) Precision: 0.308 (class 0) F1 Score: 0.366 (class 0) False Positive Rate: 13.8% True Negative Rate: 13.8% False Negative Rate: 25.5% True Positive Rate: 49.3%	allseason	VOTING CLASSIFIER: Combination of different classifiers (DecisionTree, RandomForest,K nearest neighbour,GaussNB,Extra tree,ADA boost, etc)
2	LogisticRegression	0.4850	Recall: 0.452 (class 1) Precision: 0.935 (class 1) F1 Score: 0.61 (class 1) Recall: 0.907 (class 0) Precision: 0.358 (class 0) F1 Score: 0.513 (class 0) False Positive Rate: 2.3% True Negative Rate: 2.3% True Positive Rate: 33.8% False Negative Rate: 41.0%	allseason	LOGISTIC REGRESSION:
3	VotingClassifier_knClassifier	0.4360	Recall: 0.659 (class 1) Precision: 0.789 (class 1) F1 Score: 0.718 (class 1) Recall: 0.477 (class 0) Precision: 0.32 (class 0) F1 Score: 0.383 (class 0) False Positive Rate: 13.2% True Negative Rate: 13.2% False Negative Rate: 25.5% True Positive Rate: 49.3%	allseason	K-NEAREST NEIGHBOUR: k-nearest neighbour

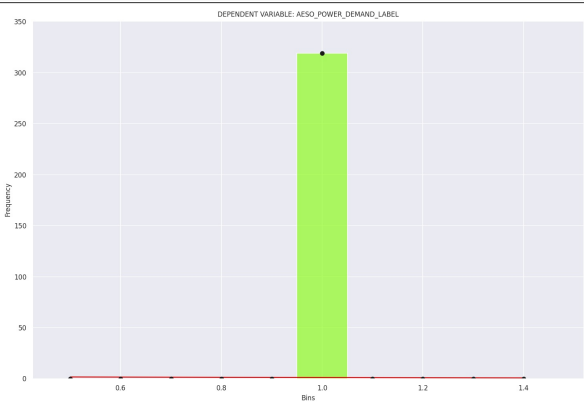
Detailed Histograms of Training and Test Data Sets

TRAINING VARIABLES

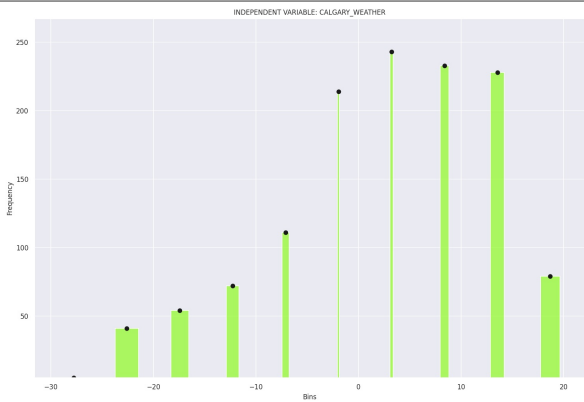


Bins	[0.00e+00, 1.00e-01]	[1.00e-01, 2.00e-01]	[2.00e-01, 3.00e-01]	[3.00e-01, 4.00e-01]	[4.00e-01, 5.00e-01]	[5.00e-01, 6.00e-01]
Count	393	0	0	0	0	0
Share	31.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Total Rows	1280	1280	1280	1280	1280	1280
Min	0.00e+00	0.00e+00	0.00e+00	0.00e+00	0.00e+00	0.00e+00
Max	1.00e+00	1.00e+00	1.00e+00	1.00e+00	1.00e+00	1.00e+00
Number of Bins	6	6	6	6	6	6

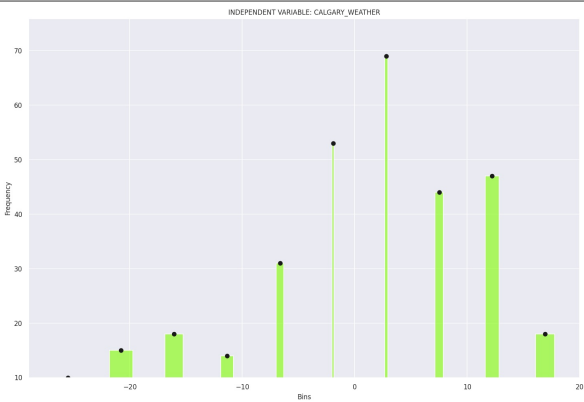
TEST VARIABLES



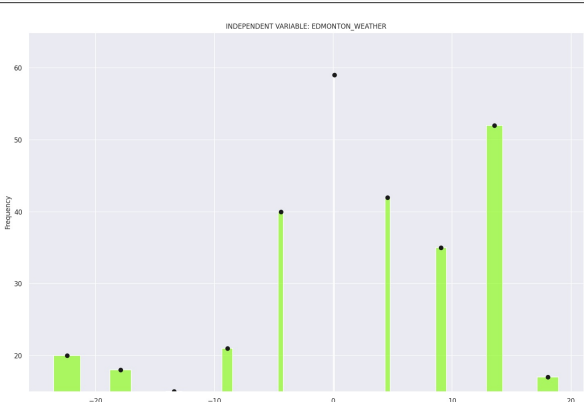
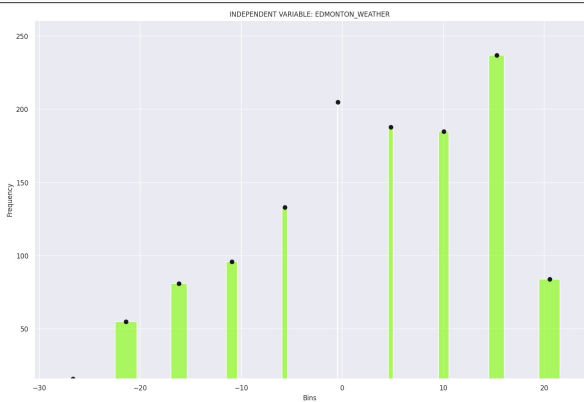
Bins	[5.00e-01, 6.00e-01]	[6.00e-01, 7.00e-01]	[7.00e-01, 8.00e-01]	[8.00e-01, 9.00e-01]	[9.00e-01, 1.00e+00]	[1.00e+00, 1.10e+00]
Count	0	0	0	0	0	319
Share	0.0%	0.0%	0.0%	0.0%	0.0%	100.0%
Total Rows	319	319	319	319	319	319
Min	1.00e+00	1.00e+00	1.00e+00	1.00e+00	1.00e+00	1.00e+00
Max	1.00e+00	1.00e+00	1.00e+00	1.00e+00	1.00e+00	1.00e+00
Number of Bins	6	6	6	6	6	6



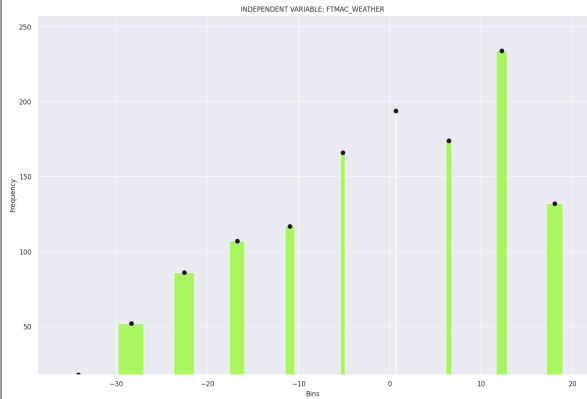
Bins	[-2.78e+01, -2.26e+01]	[-2.26e+01, -1.74e+01]	[-1.74e+01, -1.23e+01]	[-1.23e+01, -7.11e+00]	[-7.11e+00, -1.95e+00]	[-1.95e+00, 3.21e+00]
Count	5	41	54	72	111	214
Share	0.0%	3.0%	4.0%	6.0%	9.0%	17.0%
Total Rows	1280	1280	1280	1280	1280	1280
Min	-2.70e+01	-2.70e+01	-2.70e+01	-2.70e+01	-2.70e+01	-2.70e+01
Max	2.30e+01	2.30e+01	2.30e+01	2.30e+01	2.30e+01	2.30e+01
Number of Bins	6	6	6	6	6	6



Bins	[-2.55e+01, -2.08e+01]	[-2.08e+01, -1.61e+01]	[-1.61e+01, -1.14e+01]	[-1.14e+01, -6.64e+00]	[-6.64e+00, -1.93e+00]	[-1.93e+00, 2.79e+00]
Count	10	15	18	14	31	53
Share	3.0%	5.0%	6.0%	4.0%	10.0%	17.0%
Total Rows	319	319	319	319	319	319
Min	-2.50e+01	-2.50e+01	-2.50e+01	-2.50e+01	-2.50e+01	-2.50e+01
Max	2.10e+01	2.10e+01	2.10e+01	2.10e+01	2.10e+01	2.10e+01
Number of Bins	6	6	6	6	6	6

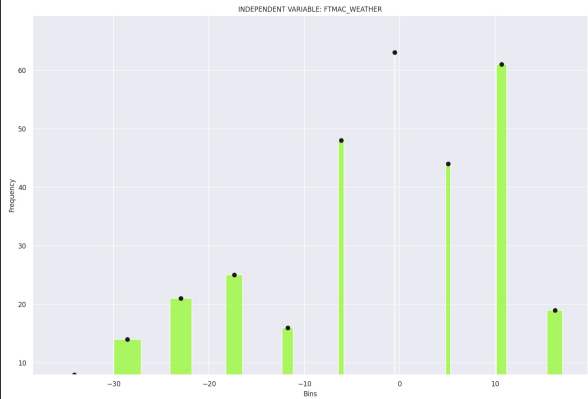


Bins	[-2.66e+01, -2.14e+01]	[-2.14e+01, -1.62e+01]	[-1.62e+01, -1.09e+01]	[-1.09e+01, -5.68e+00]	[-5.68e+00, -4.45e-01]	[-4.45e-01, 4.79e+00]
Count	16	55	81	96	133	205
Share	1.0%	4.0%	6.0%	8.0%	10.0%	16.0%
Total Rows	1280	1280	1280	1280	1280	1280
Min	-2.60e+01	-2.60e+01	-2.60e+01	-2.60e+01	-2.60e+01	-2.60e+01
Max	2.50e+01	2.50e+01	2.50e+01	2.50e+01	2.50e+01	2.50e+01
Number of Bins	6	6	6	6	6	6



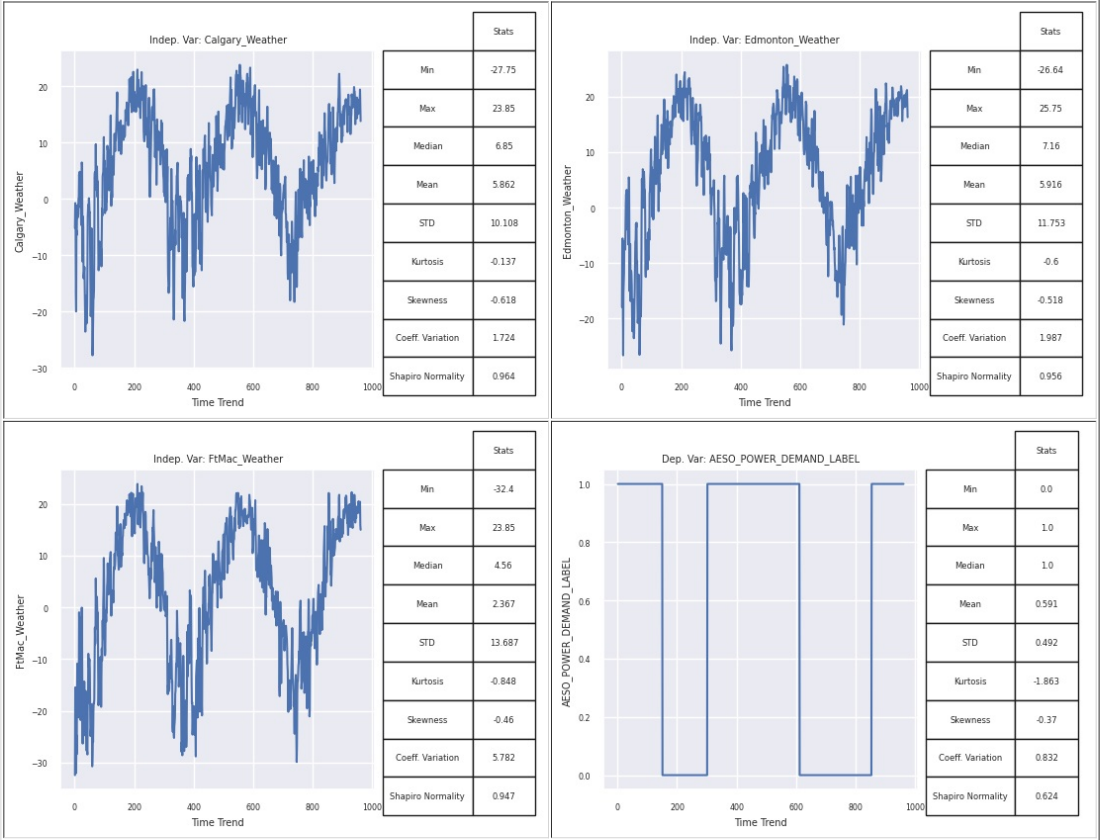
Bins	[-3.41e+01, -2.83e+01]	[-2.83e+01, -2.25e+01]	[-2.25e+01, -1.68e+01]	[-1.68e+01, -1.09e+01]	[-1.09e+01, -5.15e+00]	[-5.15e+00, 6.50e-01]
Count	18	52	86	107	117	166
Share	1.0%	4.0%	7.0%	8.0%	9.0%	13.0%
Total Rows	1280	1280	1280	1280	1280	1280
Min	-3.40e+01	-3.40e+01	-3.40e+01	-3.40e+01	-3.40e+01	-3.40e+01
Max	2.30e+01	2.30e+01	2.30e+01	2.30e+01	2.30e+01	2.30e+01
Number of Bins	6	6	6	6	6	6

Bins	[-2.24e+01, -1.79e+01]	[-1.79e+01, -1.34e+01]	[-1.34e+01, -8.91e+00]	[-8.91e+00, -4.42e+00]	[-4.42e+00, 7.50e-02]	[7.50e-02, 4.57e+00]
Count	20	18	15	21	40	59
Share	6.0%	6.0%	5.0%	7.0%	13.0%	18.0%
Total Rows	319	319	319	319	319	319
Min	-2.20e+01	-2.20e+01	-2.20e+01	-2.20e+01	-2.20e+01	-2.20e+01
Max	2.20e+01	2.20e+01	2.20e+01	2.20e+01	2.20e+01	2.20e+01
Number of Bins	6	6	6	6	6	6



Bins	[-3.41e+01, -2.85e+01]	[-2.85e+01, -2.29e+01]	[-2.29e+01, -1.74e+01]	[-1.74e+01, -1.18e+01]	[-1.18e+01, -6.15e+00]	[-6.15e+00, -5.50e-01]
Count	8	14	21	25	16	48
Share	3.0%	4.0%	7.0%	8.0%	5.0%	15.0%
Total Rows	319	319	319	319	319	319
Min	-3.40e+01	-3.40e+01	-3.40e+01	-3.40e+01	-3.40e+01	-3.40e+01
Max	2.10e+01	2.10e+01	2.10e+01	2.10e+01	2.10e+01	2.10e+01
Number of Bins	6	6	6	6	6	6

Detailed Graphs of Variables Against Time



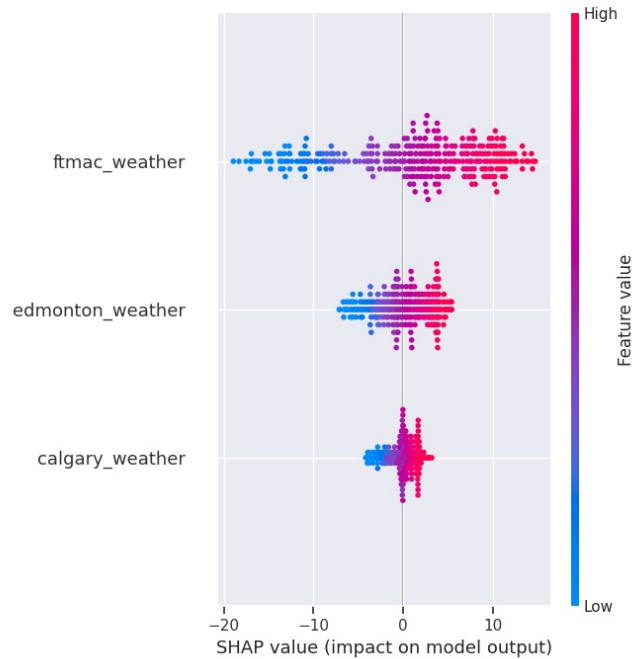
Confusion Matrix

	Predicted 0	Predicted 1
True 0	68	83
True 1	1.5e+02	3e+02

True Positives: 83
False Positives: 68
True Negatives: 153
False Negatives: 153
Total Population: 600
The False Positive Rate(FPR) is: 13.83%
The False Negative Rate is: 25.5%
The True Positive Rate is: 49.33%
The True Negative Rate is: 11.33%
The Positive Likelihood Ratio (True Positive Rate/False Positive Rate)is: 3.57
The Negative Likelihood Ratio (False Negative Rate/True Negative Rate) is: 2.25
Accuracy: 0.5
Precision: 0.781
Recall: 0.659
F1 Score: 0.715

Recall Curve: [1.000, 0.998, 0.989, 0.973, 0.971, 0.955, 0.935, 0.920, 0.902, 0.891, 0.889, 0.886, 0.880, 0.862, 0.855, 0.849, 0.844, 0.837, 0.835, 0.824, 0.817, 0.811, 0.806, 0.804, 0.804, 0.800, 0.797, 0.793, 0.793, 0.791, 0.791, 0.784, 0.784, 0.777, 0.775, 0.768, 0.759, 0.757, 0.757, 0.757, 0.751, 0.746, 0.744, 0.744, 0.744, 0.742, 0.739, 0.737, 0.735, 0.731, 0.731, 0.728, 0.726, 0.724, 0.722, 0.719, 0.717, 0.717, 0.713, 0.710, 0.704, 0.702, 0.699, 0.693, 0.688, 0.686, 0.684, 0.675, 0.673, 0.670, 0.668, 0.664, 0.661, 0.659, 0.657, 0.648, 0.637, 0.619, 0.606, 0.588, 0.577, 0.559, 0.548, 0.530, 0.519, 0.506, 0.492, 0.490, 0.483, 0.474, 0.465, 0.454, 0.439, 0.430, 0.419, 0.416, 0.401, 0.396, 0.394, 0.392, 0.388, 0.376, 0.372, 0.367, 0.363, 0.361, 0.356, 0.350, 0.345, 0.345, 0.334, 0.323, 0.316, 0.310, 0.303, 0.296, 0.290, 0.281, 0.274, 0.269, 0.261, 0.258, 0.256, 0.254, 0.245, 0.238, 0.232, 0.229, 0.227, 0.223, 0.218, 0.214, 0.209, 0.200, 0.185, 0.183, 0.180, 0.176, 0.174, 0.171, 0.163, 0.158, 0.151, 0.145, 0.138, 0.129, 0.109, 0.096, 0.087, 0.076, 0.071, 0.067, 0.060, 0.056, 0.053, 0.045, 0.042, 0.040, 0.036, 0.033, 0.029, 0.024, 0.022, 0.020, 0.018, 0.013, 0.009, 0.007, 0.004, 0.002, 0.000]

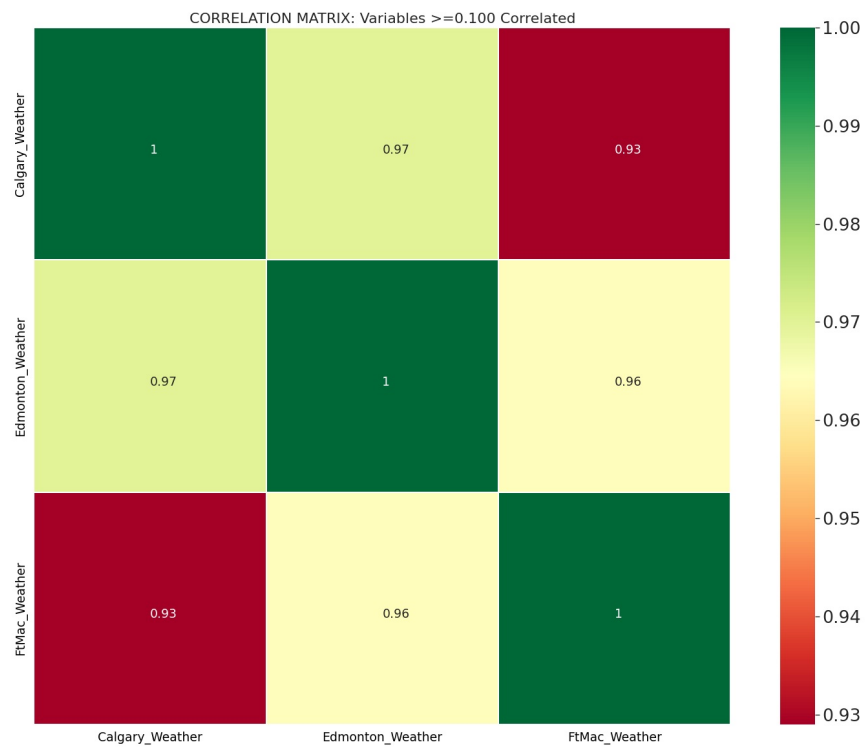
MODEL EXPLANATION



- The x-axis represents the model's output values of **AESO_POWER_DEMAND_LABEL**.
- The plot is centered on the x-axis at `explainer.expected_value`.
- All values are relative to the model's expected value like a linear model's effects are relative to the intercept.
- The y-axis lists the model's features. By default, the features are ordered by descending importance.
- The importance is calculated over the observations plotted. This is usually different than the importance ordering for the entire dataset.
- In addition to feature importance ordering, the decision plot also supports hierarchical cluster feature ordering and user-defined feature ordering.
- Each observation's prediction is represented by a colored line.
- At the top of the plot, each line strikes the x-axis at its corresponding observation's predicted value. This value determines the color of the line on a spectrum.
- Moving from the bottom of the plot to the top, SHAP values for each feature are added to the model's base value.
- This shows how each feature contributes to the overall prediction.
- At the bottom of the plot, the observations converge at `explainer.expected_value`.
- The points in the graph are the values of the feature in the training dataset.

FEATURE SELECTION	
RFE Variable (Most important to Least Important)	Value
FtMac_Weather	0.008
Edmonton_Weather	0.005
Calgary_Weather	0.003
Best Variable(s) From Genetic Algorithm	
Calgary_Weather	
Excluded Variable(s)	
FtMac_Weather	
Edmonton_Weather	
PCA for Best Variable(s)	Value
AESO_POWER_DEMAND_LABEL_pca_1	0.707
AESO_POWER_DEMAND_LABEL_pca_2	-0.707
Calgary_Weather_pca_1	-0.707
Calgary_Weather_pca_2	-0.707
PCA Explained Variance	Value
PCA1	0.551
PCA2	0.449
<ul style="list-style-type: none">• Feature selection shows which variables were more influential than other variables• It uses two core algorithms: Recursive Feature Elimination (RFE) and Genetic Algorithm to determine influence• It also performs PCA (principal component analysis) analysis to determine the influence of the best variables in the model• These results should be used in conjunction with other information as well as theory to establish relevance and confidence in the chosen model formulation	

CORRELATION MATRIX



CORRELATED FEATURES			
	Feature(s)	Feature(s)	Correlation >= 0.100
0	Calgary_Weather	FtMac_Weather	0.929
1	Edmonton_Weather	FtMac_Weather	0.964
2	Calgary_Weather	Edmonton_Weather	0.970
3	Calgary_Weather	Calgary_Weather	NaN

SUGGESTED CORRELATED FEATURES TO DELETE		
	2 Feature(s) to Delete	Correlation
0	Calgary_Weather	0.929
1	Edmonton_Weather	0.964

END OF REPORT

MAADSBML Python Library: <https://pypi.org/project/maadsbml/>
MAADSBML Docker Container For Windows: <https://hub.docker.com/r/maadsdocker/maads-batch-automl-otics>
MAADSBML Docker Container For MAC: <https://hub.docker.com/r/maadsdocker/maads-batch-automl-otics-arm64>
MAADSBML Sample Code and Setup: <https://github.com/smaurice101/raspberrypi/tree/main/maadsbml>

MAADSBML
Developed and Maintained by: Otics Advanced Analytics, Inc.
Toronto, Ontario, Canada
<https://www.otics.ca>
Email: support@otics.ca