

Radiogenomics Analysis Report

09/08/2021 23:11:32

-----Model inputs-----

Mode:Train

Model:LinearRegression

Params:default

No. of imaging features provided: 540

No. of gene features provided:976

SampleID check results: 'The SampleIDs match for imaging and gene features'

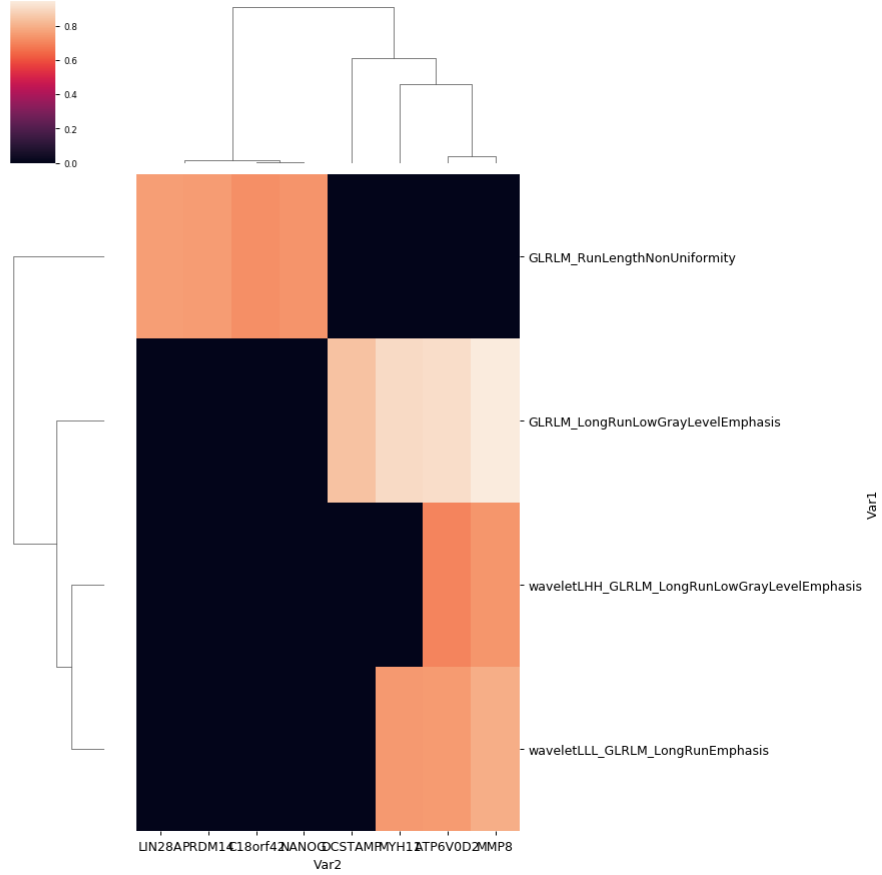
No. of samples: 106

performing Stand_scaler normalization for imaging features

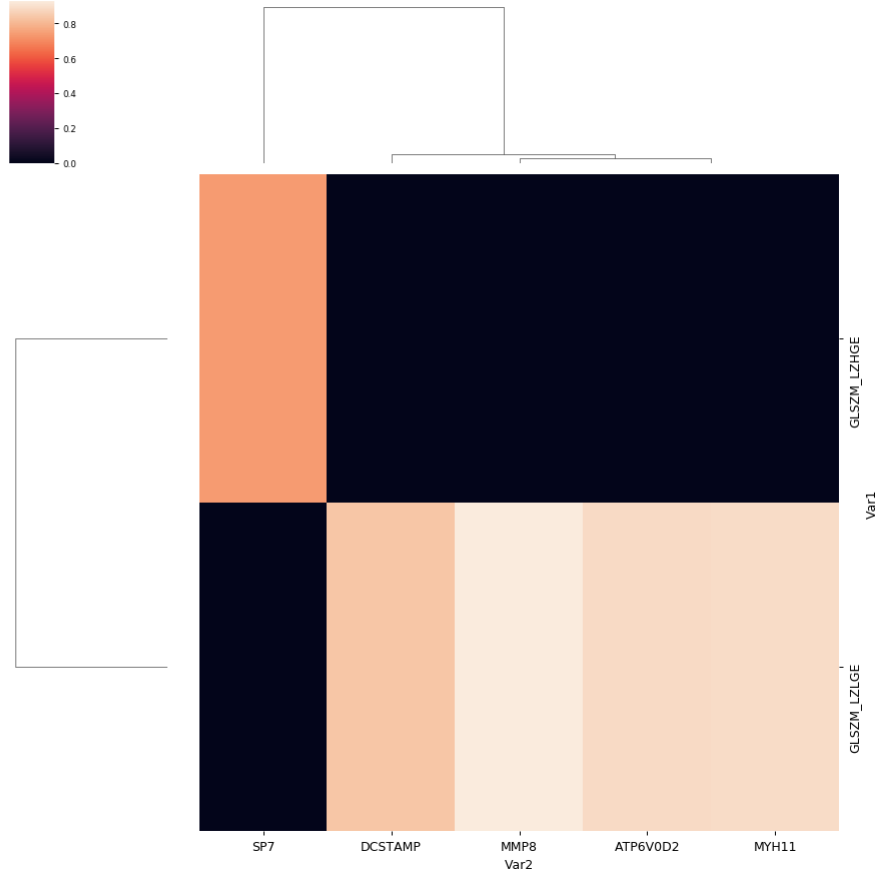
performing Stand_scaler for gene features

-----Multivariate Correlations (pearson based)-----

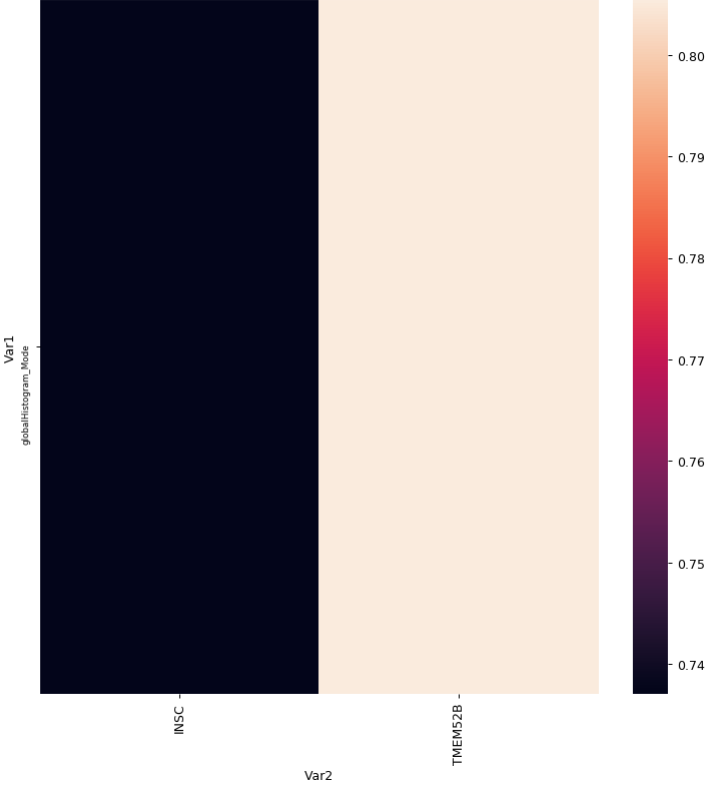
Top significant correlations (FDR_adjusted_pValue<0.05) for GLRLM features



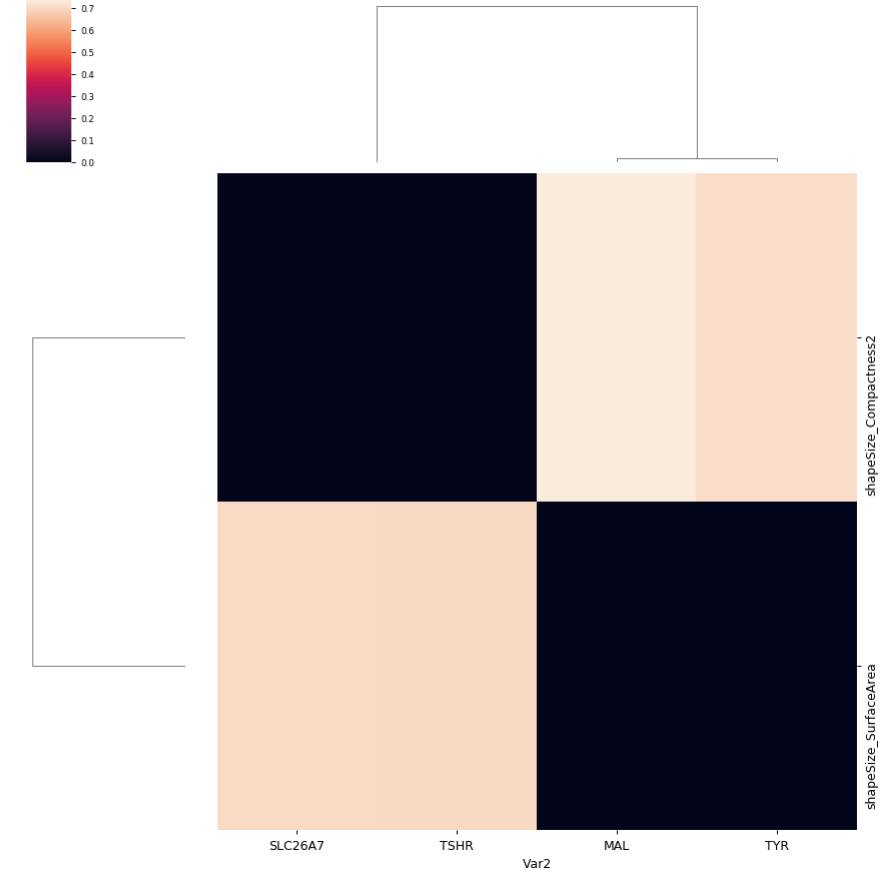
Top significant correlations (FDR_adjusted_pValue<0.05) for GLSZM features



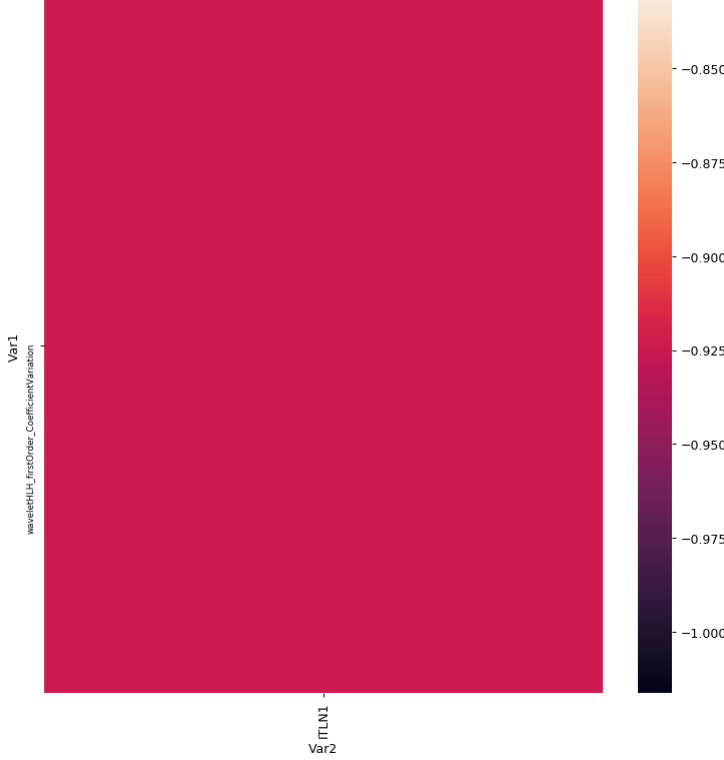
Top significant correlations (FDR_adjusted_pValue<0.05) for globalHistogram features



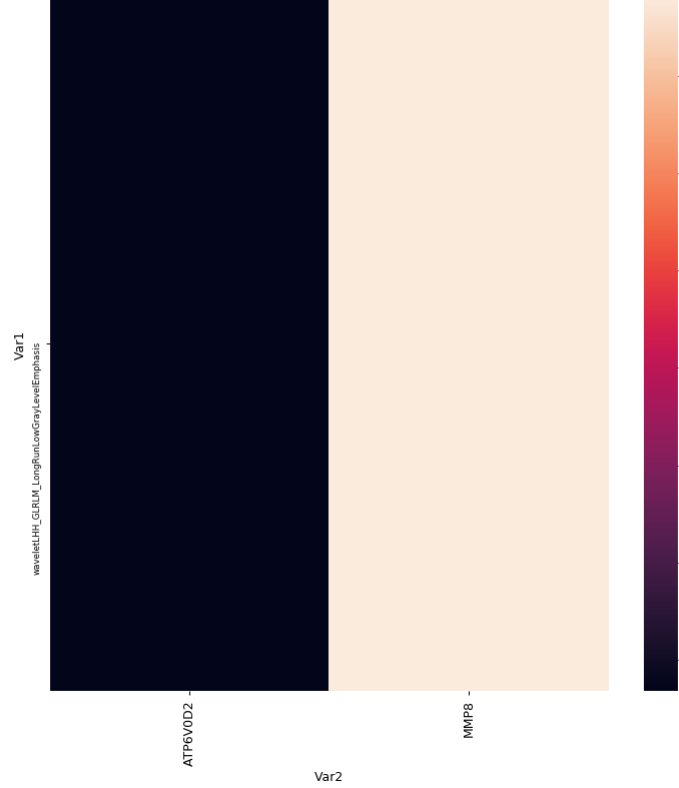
Top significant correlations (FDR_adjusted_pValue<0.05) for shapeSize features



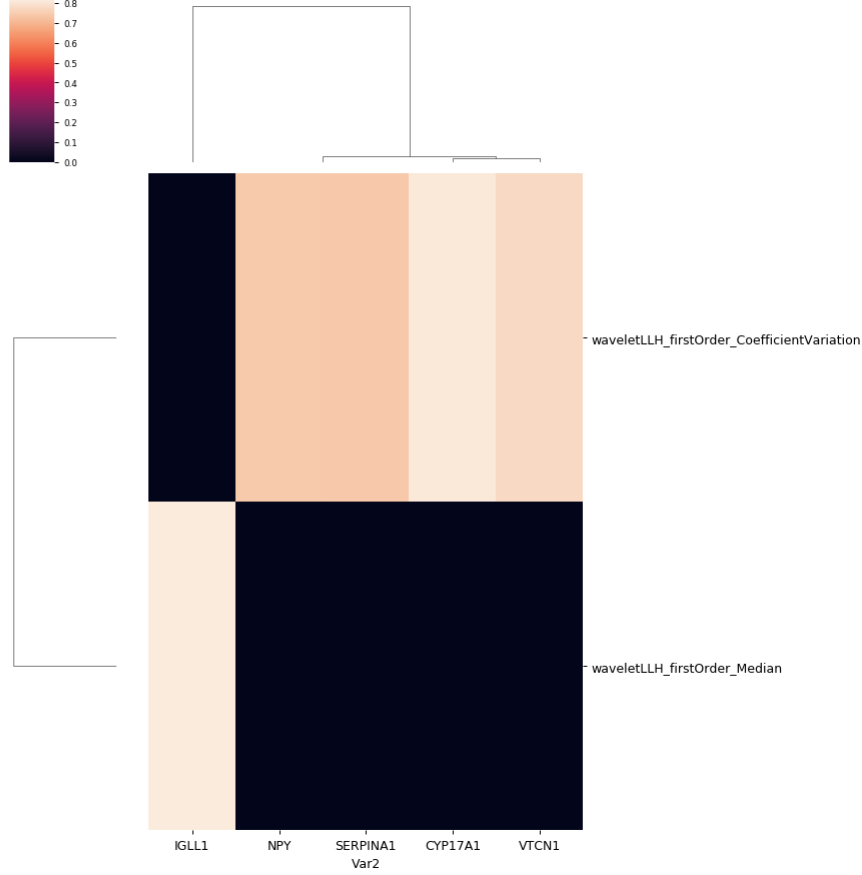
Top significant correlations (FDR_adjusted_pValue<0.05) for waveletHLH features



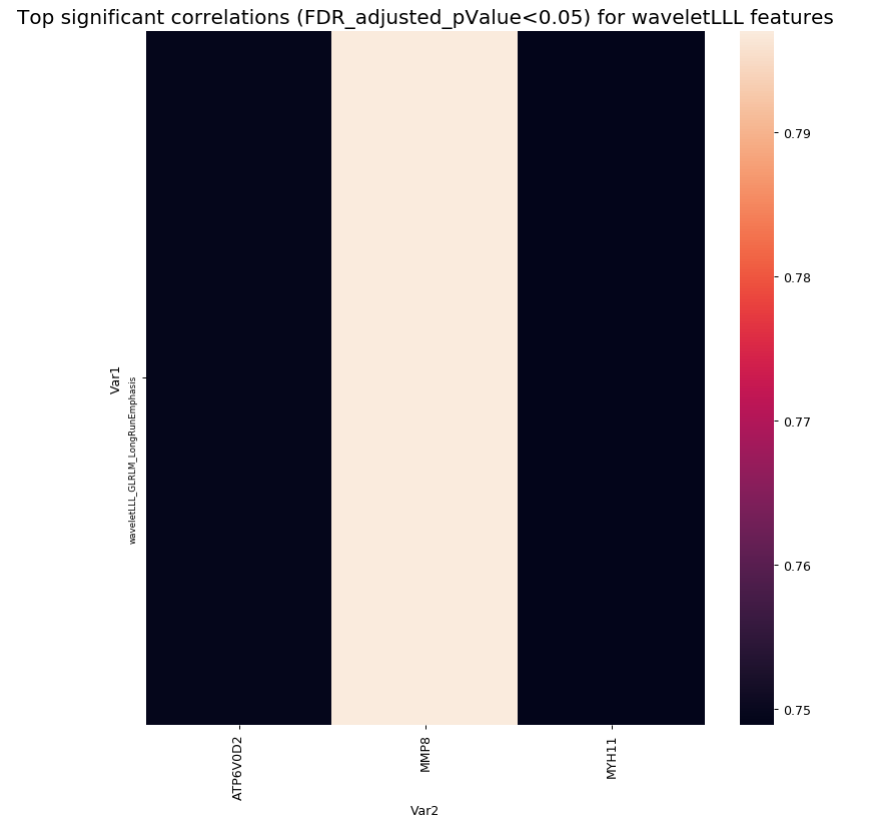
Top significant correlations (FDR_adjusted_pValue<0.05) for waveletLHH features



Top significant correlations (FDR_adjusted_pValue<0.05) for waveletLLH features



Top significant correlations (FDR_adjusted_pValue<0.05) for waveletLLL features



-----Features with highly significant correlations-----

Below is the list of imaging features

['GLRLM_LongRunLowGrayLevelEmphasis', 'GLRLM_RunLengthNonUniformity', 'GLSZM_LZHGE', 'GLSZM_LZLGE', 'globalHistogram_Mode', 'shapeSize_Compactness2', 'shapeSize_SurfaceArea', 'waveletLH_firstOrder_CoefficientVariation', 'waveletLH_GLRM_LongRunLowGrayLevelEmphasis', 'waveletLH_firstOrder_CoefficientVariation', 'waveletLH_firstOrder_Median', 'waveletLLL_GLRM_LongRunEmphasis']

Below is the list of gene features

['ATP6V0D2', 'C18orf42', 'CYP17A1', 'DCSTAMP', 'IGLL1', 'INSC', 'ITLN1', 'LIN28A', 'MAL', 'MMP8', 'MYH11', 'NANOG', 'NPY', 'PRDM14', 'SERPINA1', 'SLC26A7', 'SP7', 'TMEM52B', 'TSHR', 'TYR', 'VTCN1']

Number of imaging features

12

Number of gene features

21

-----Number of Samples for Training and Testing-----

No. of samples for training:84

No. of samples for test:22

-----Model Summary-----

Model Type : LinearRegression

Cross Validation Metrics:

Parameters: cv=3 scoring=neg_mean_squared_error

Cross validation score:2.16718355799

Model Parameters:

copy_X:True

normalize:False

n_jobs:None

fit_intercept:True

Min Square Error for the Model

MSE of train_eval set:0.302017441903

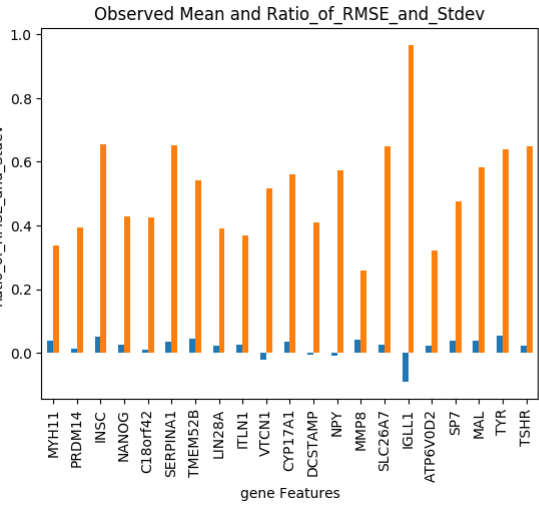
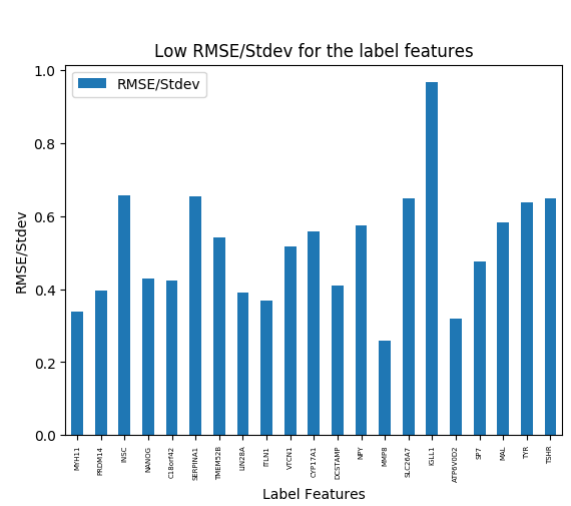
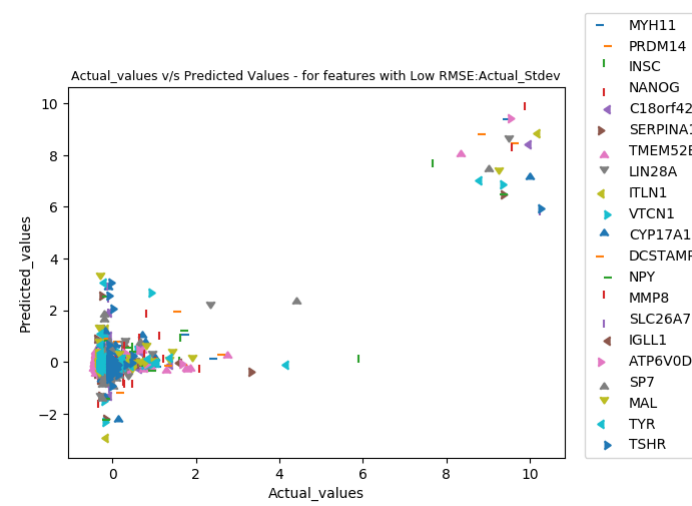
No. of features showing LOW 'RMSE/Stdev' (≤ 1.0): 21

All such features with their Low 'RMSE/Stdev' values could be found in output file: train_eval_LinearRegression_Labels_with_Low_Ratio.csv

No. of features showing HIGH 'RMSE/Stdev' (> 1.0): 0

All such features with their High 'RMSE/Stdev' values could be found in output file: train_eval_LinearRegression_Labels_with_High_Ratio.csv

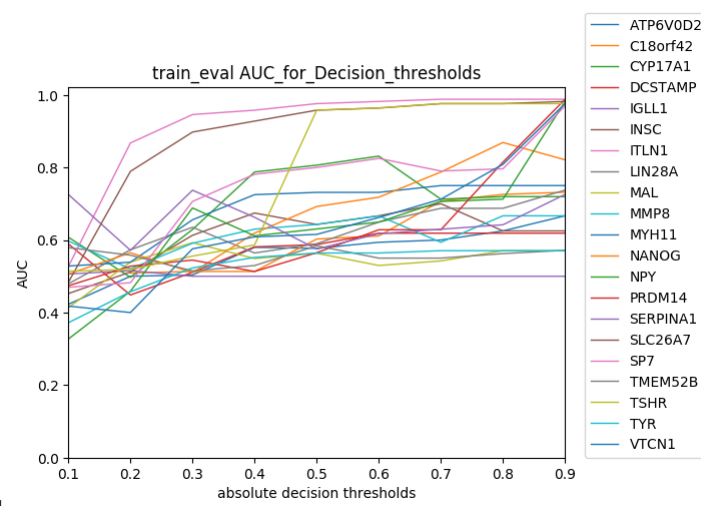
Model evaluation for Train data for label features showing Low 'RMSE/Stdev' (≤ 1.0)



Observed Mean

Ratio_of_RMSE_and_Stdev

Content-type: text/html



Content-type: text/html

-----Model evaluation for Test data-----

Min Square Error for the Model

MSE of test_eval set:0.572402014754

No. of features showing LOW 'RMSE/Stdev' (≤ 1.0): 1

All such features with their Low 'RMSE/Stdev' values could be found in output file: test_eval_LinearRegression_Labels_with_Low_Ratio.csv

No. of features showing HIGH 'RMSE/Stdev' (> 1.0): 20

All such features with their High 'RMSE/Stdev' values could be found in output file: test_eval_LinearRegression_Labels_with_High_Ratio.csv

Model evaluation for Test data for label features showing Low 'RMSE/Stdev' (≤ 1.0)

