Radiogenomics Analysis Report

22/09/2021 16:58:55

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

Mode:Train

Model:multiTaskLASSO

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'
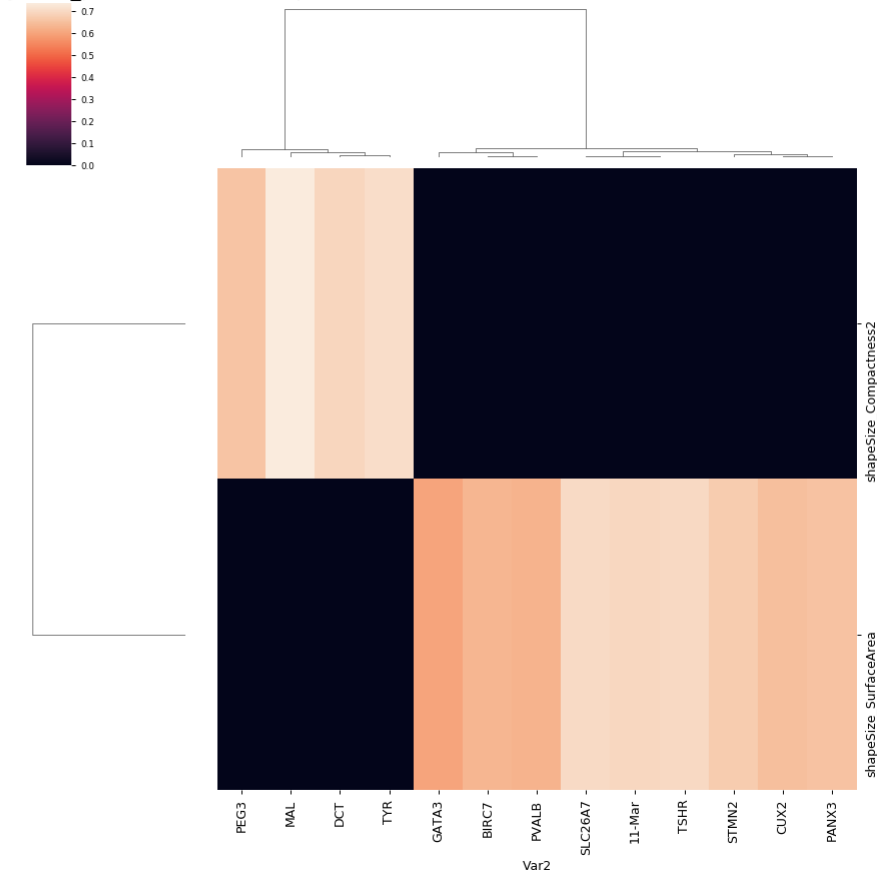
No. of samples: 106

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
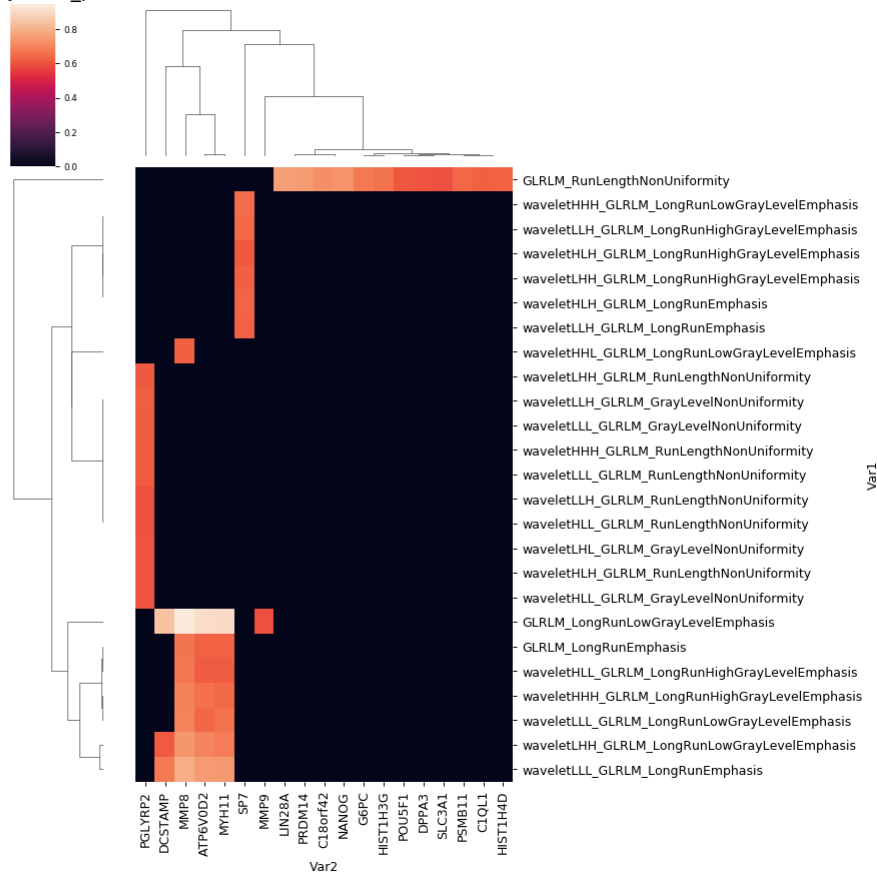
Top significant correlations (FDR_adjusted_pValue<0.05) for shapeSize features
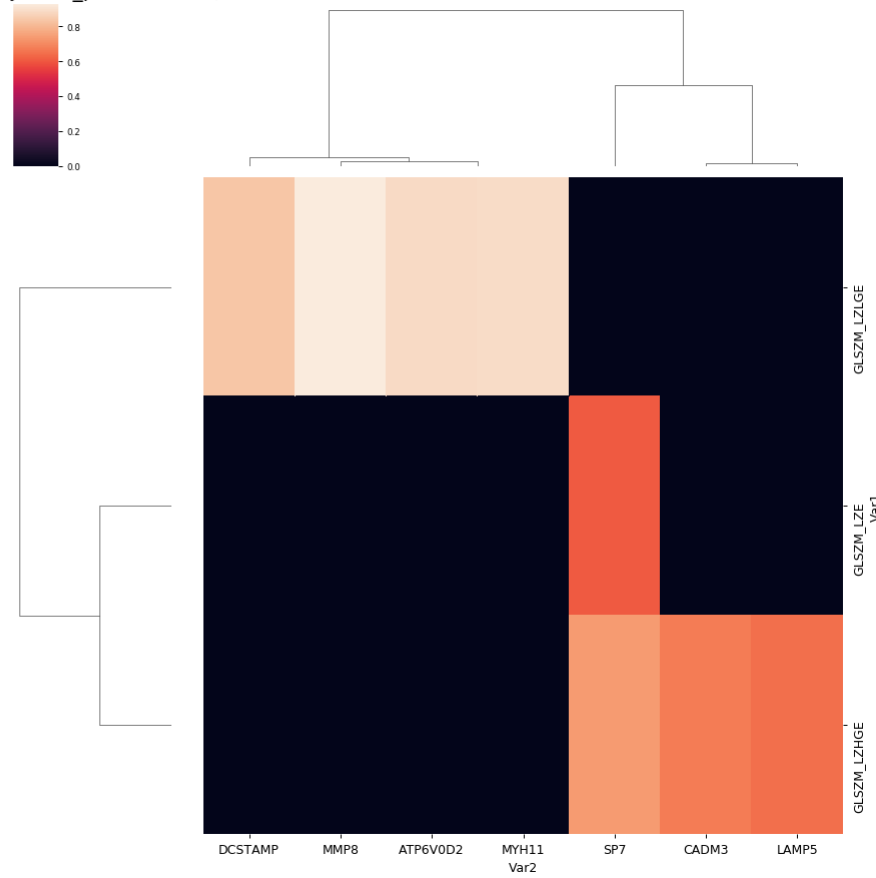
Top significant correlations (FDR_adjusted_pValue<0.05) for waveletHHH features
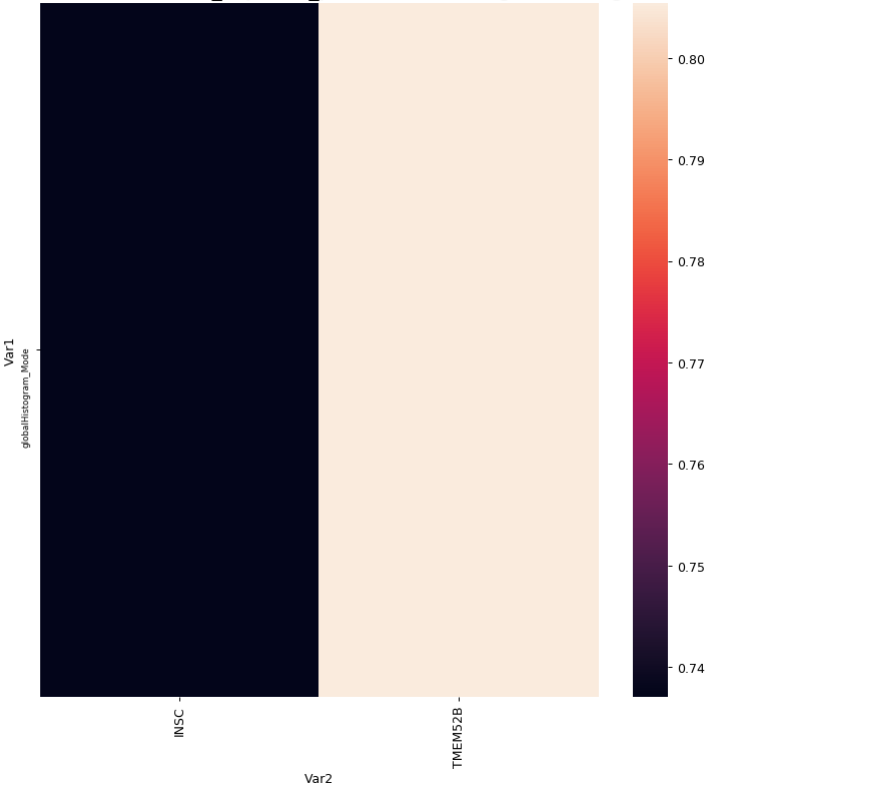
Top significant correlations (FDR_adjusted_pValue<0.05) for waveletHHL features

Top significant correlations (FDR_adjusted_pValue<0.05) for waveletHLH features

Top significant correlations (FDR_adjusted_pValue<0.05) for waveletHLL features

Top significant correlations (FDR_adjusted_pValue<0.05) for waveletLHH features

Top significant correlations (FDR_adjusted_pValue<0.05) for waveletLHL features
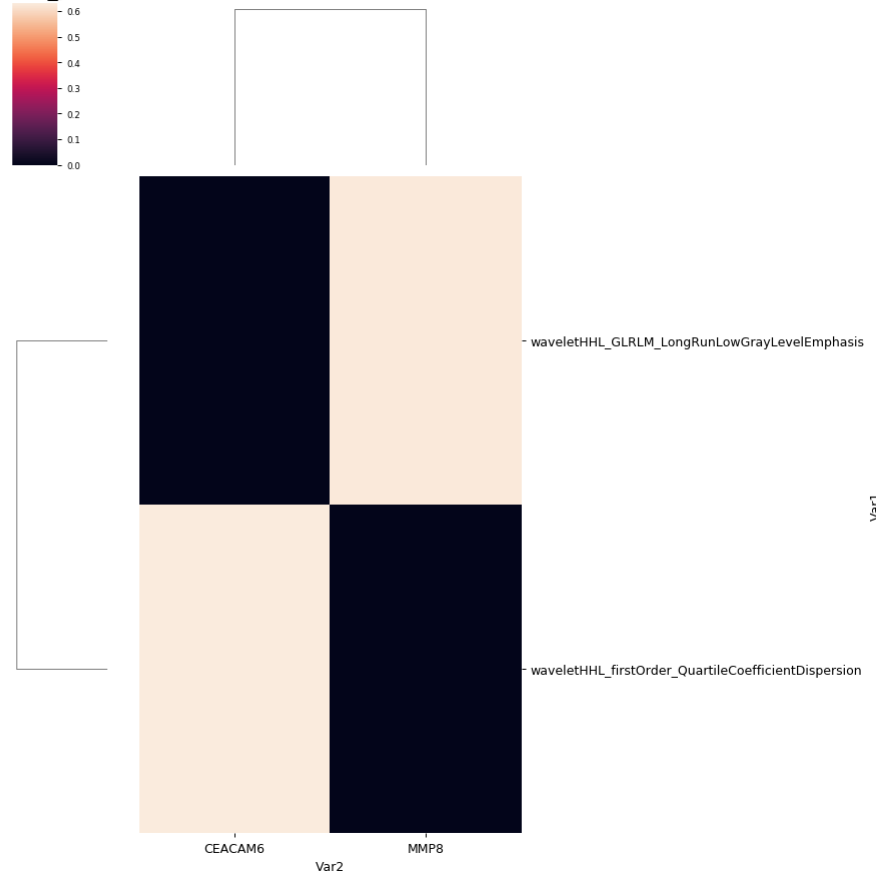
Top significant correlations (FDR_adjusted_pValue<0.05) for waveletLLH features

Top significant correlations (FDR_adjusted_pValue<0.05) for waveletLLL features
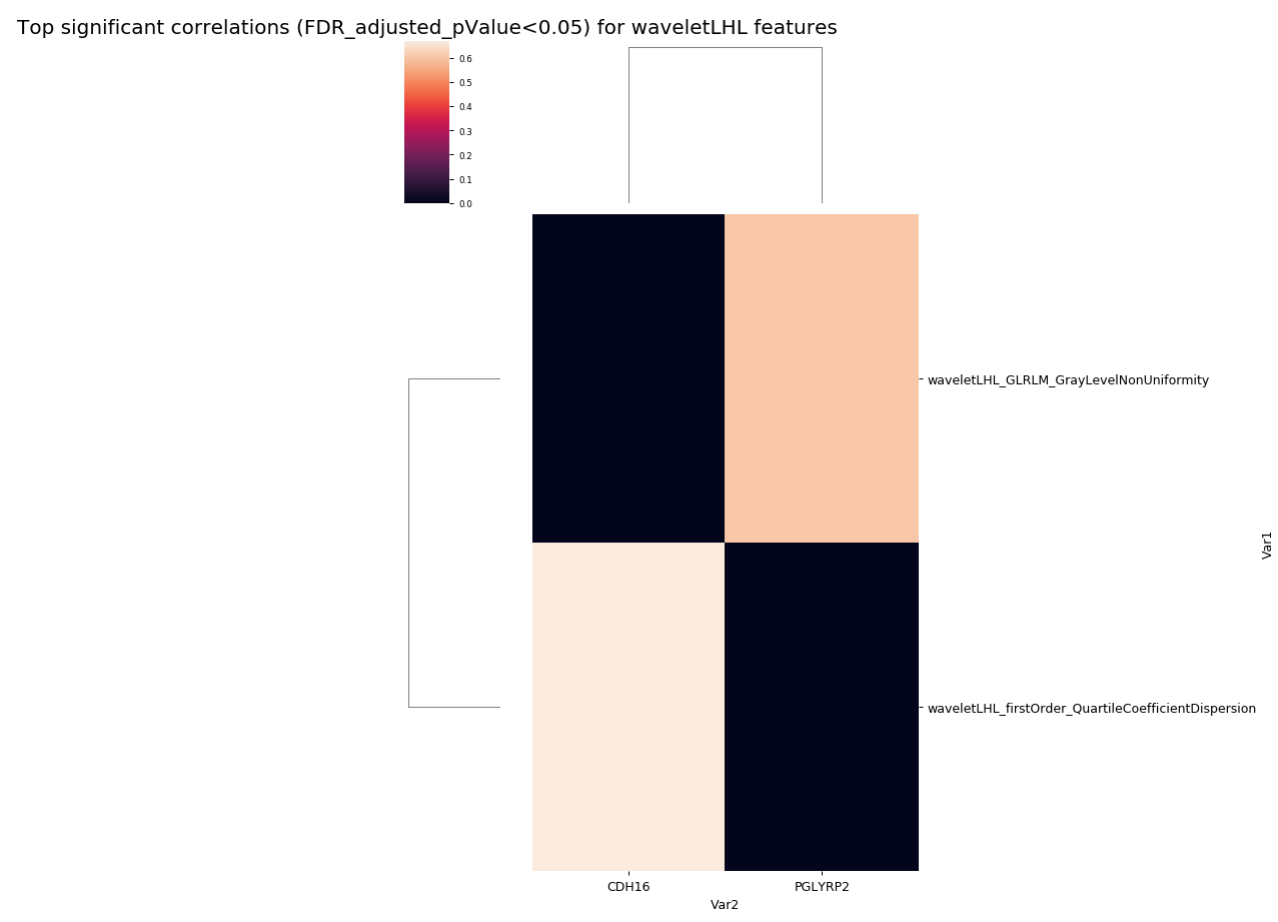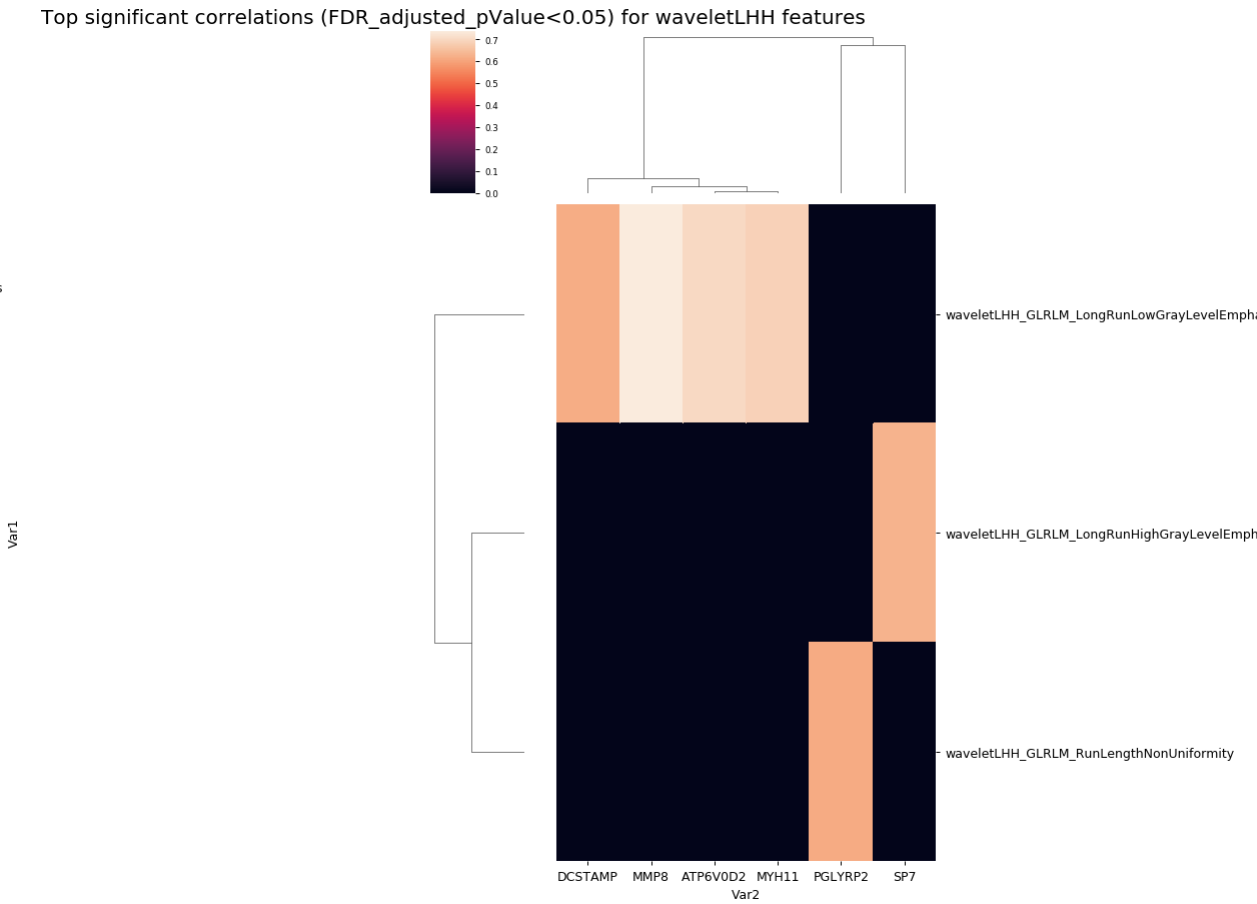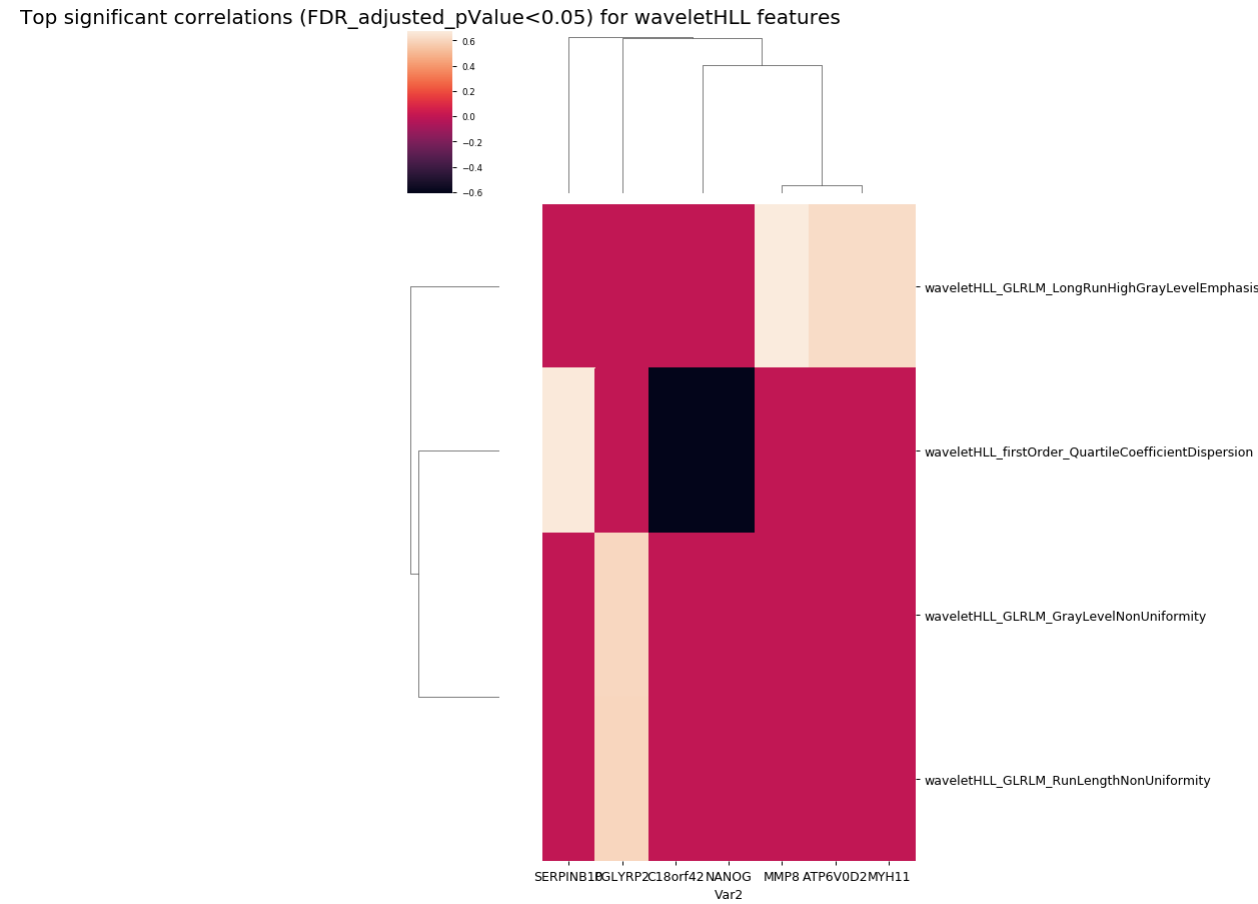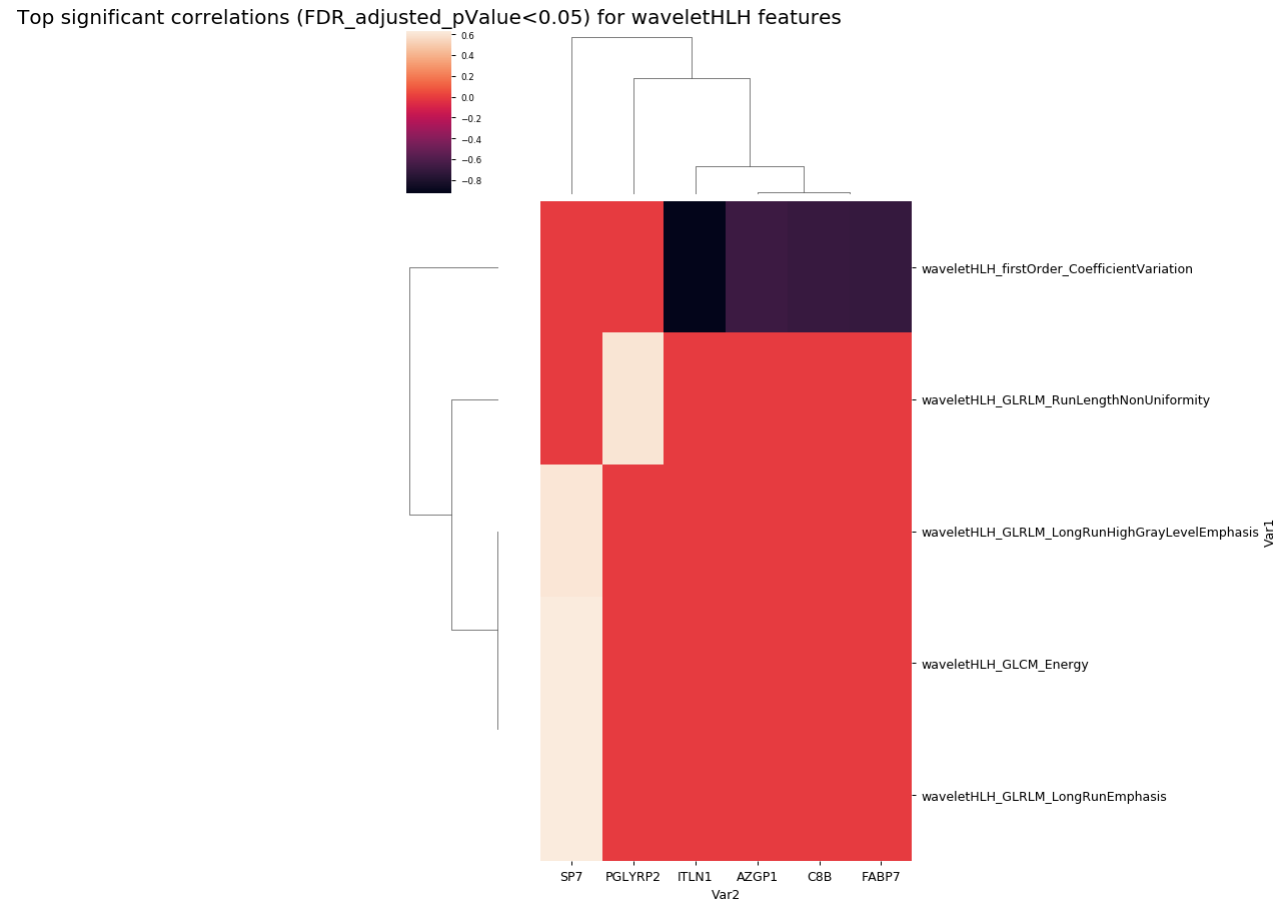
--------------------------------Features with highly significant correlations--------------------------------

Below is the list of imaging features

['GLRLM_LongRunEmphasis', 'GLRLM_LongRunLowGrayLevelEmphasis', 'GLRLM_RunLengthNonUniformity', 'GLSZM_LZE', 'GLSZM_LZHGE', 'GLSZM_LZLGE', 'globalHistogram_Mode', 'shapeSize_Compactness2', 'shapeSize_SurfaceArea', 'waveletHHH_GLRLM_LongRunHighGrayLevelEmphasis', 'waveletHHH_GLRLM_LongRunLowGrayLevelEmphasis', 'waveletHHH_GLRLM_RunLengthNonUniformity', 'waveletHHL_GLRLM_LongRunLowGrayLevelEmphasis', 'waveletHHL_firstOrder_QuartileCoefficientDispersion', 'waveletHLH_GLCM_Energy', 'waveletHLH_GLRLM_LongRunEmphasis', 'waveletHLH_GLRLM_LongRunHighGrayLevelEmphasis', 'waveletHLH_GLRLM_RunLengthNonUniformity', 'waveletHLH_firstOrder_CoefficientVariation', 'waveletHLL_GLRLM_GrayLevelNonUniformity', 'waveletHLL_GLRLM_LongRunHighGrayLevelEmphasis', 'waveletHLL_GLRLM_RunLengthNonUniformity', 'waveletHLL_firstOrder_QuartileCoefficientDispersion', 'waveletLHH_GLRLM_GrayLevelNonUniformity', 'waveletLHH_GLRLM_LongRunHighGrayLevelEmphasis', 'waveletLHH_GLRLM_LongRunLowGrayLevelEmphasis', 'waveletLHH_GLRLM_RunLengthNonUniformity', 'waveletLHL_GLRLM_GrayLevelNonUniformity', 'waveletLHL_firstOrder_QuartileCoefficientDispersion', 'waveletLLH_GLRLM_GrayLevelNonUniformity', 'waveletLLH_GLRLM_LongRunEmphasis', 'waveletLLH_GLRLM_LongRunHighGrayLevelEmphasis', 'waveletLLH_GLRLM_RunLengthNonUniformity', 'waveletLLH_firstOrder_CoefficientVariation', 'waveletLLH_firstOrder_Mean', 'waveletLLH_firstOrder_Median', 'waveletLLL_GLCM_Energy', 'waveletLLL_GLRLM_GrayLevelNonUniformity', 'waveletLLL_GLRLM_LongRunEmphasis', 'waveletLLL_GLRLM_LongRunLowGrayLevelEmphasis', 'waveletLLL_GLRLM_RunLengthNonUniformity', 'waveletLLL_firstOrder_Maximum']

Below is the list of gene features

['11-Mar', 'ADRA2C', 'AGXT', 'ATP6V0D2', 'AZGP1', 'BIRC7', 'C18orf42', 'C1QL1', 'C8B', 'CADM3', 'CDH16', 'CEACAM6', 'CUX2', 'CYP17A1', 'DCSTAMP', 'DCT', 'DPPA3', 'EEF1A2', 'FABP7', 'G6PC', 'GATA3', 'GFRA3', 'HIST1H3G', 'HIST1H4D', 'IGLL1', 'INSC', 'ITLN1', 'KLK4', 'LAMP5', 'LIN28A', 'MAL', 'MMP8', 'MMP9', 'MYH11', 'NANOG', 'NPY', 'PANX3', 'PEG3', 'PGLYRP2', 'POU5F1', 'PRDM14', 'PSMB11', 'PVALB', 'SERPINA1', 'SERPINB10', 'SLC22A3', 'SLC26A7', 'SLC3A1', 'SP7', 'STMN2', 'TMEM52B', 'TSHR', 'TYR', 'VTCN1']

Number of imaging features

41

Number of gene features

54

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

No. of samples for training:84

No. of samples for test:22

--------------------------**Model Summary**--------------------------

Model Type : multiTaskLASSO

Cross Validation Metrics:

Parameters: cv=3 scoring:neg_mean_squared_error

Cross validation score:0.9291072857728

Model Parameters:

normalize:False

warm_start:False

selection:cyclic

fit_intercept:True

max_iter:1000

random_state:None

tol:0.0001

copy_X:True

alpha:1.0

----------------**Model evaluation for Train data**----------------

Min Square Error for the Model

MSE of train_eval set:0.56114653477

No. of features showing LOW 'RMSE/Stdev' values (<=1.0): 54

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

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

Content-type: text/html   Content-type: text/html

---------------**Model evaluation for Test data**--------------------
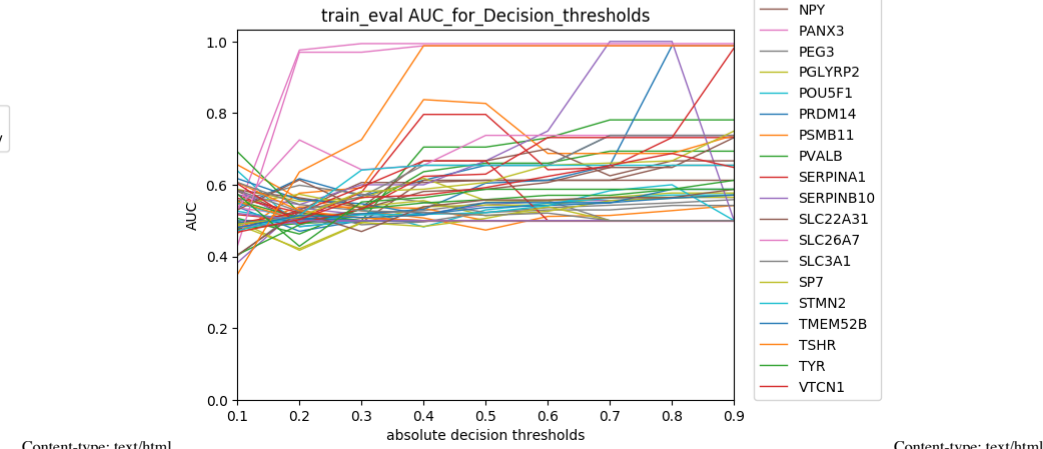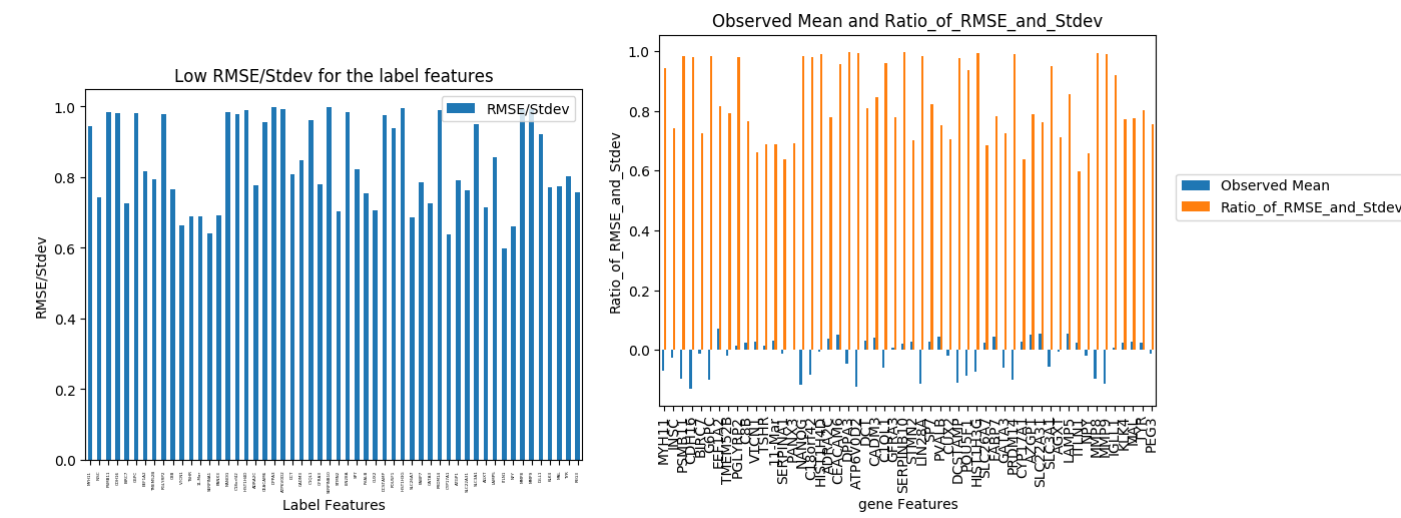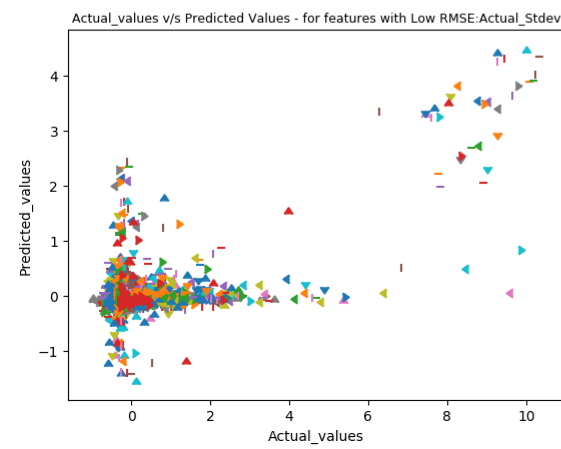
Min Square Error for the Model

MSE of test_eval set:1.5231959094

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_multiTaskLASSO_Labels_with_Low_Ratio.csv

No. of features showing HIGH 'RMSE/Stdev' (>1.0): 53

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

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

Content-type: text/html   Content-type: text/html