

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

Mode:Train

Model:DecisionTree

Params:default

No. of imaging features provided: 540

No. of gene features provided:976

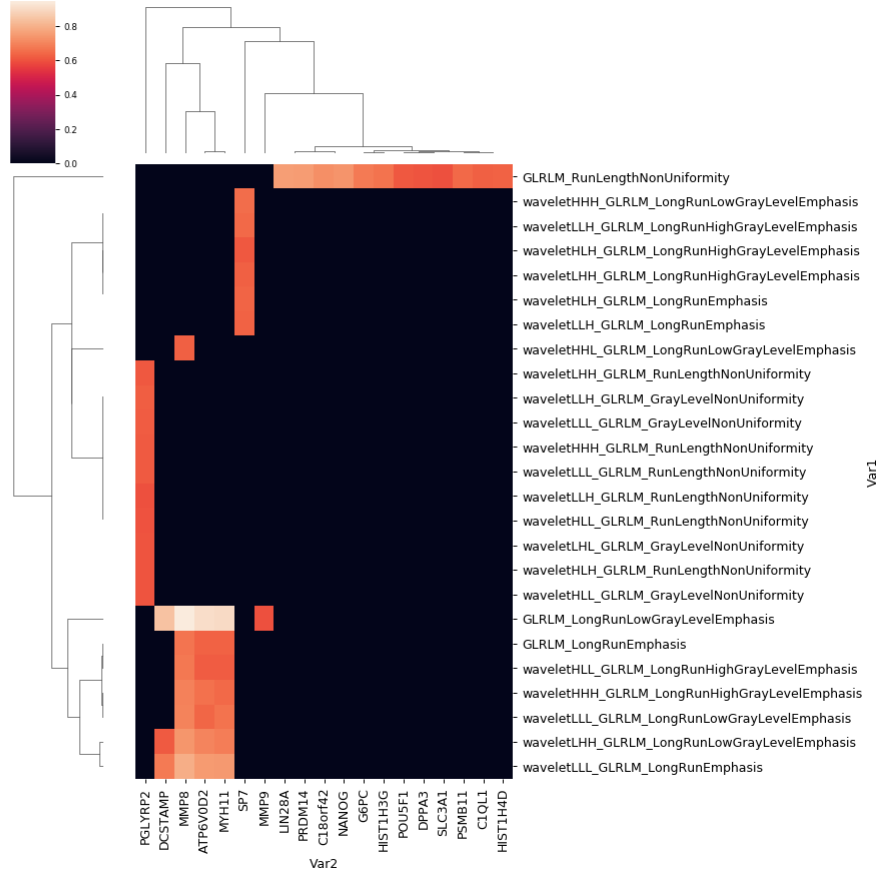
SampleID check results: 'The SampleIDs match for imaging and gene features'

No. of samples: 106

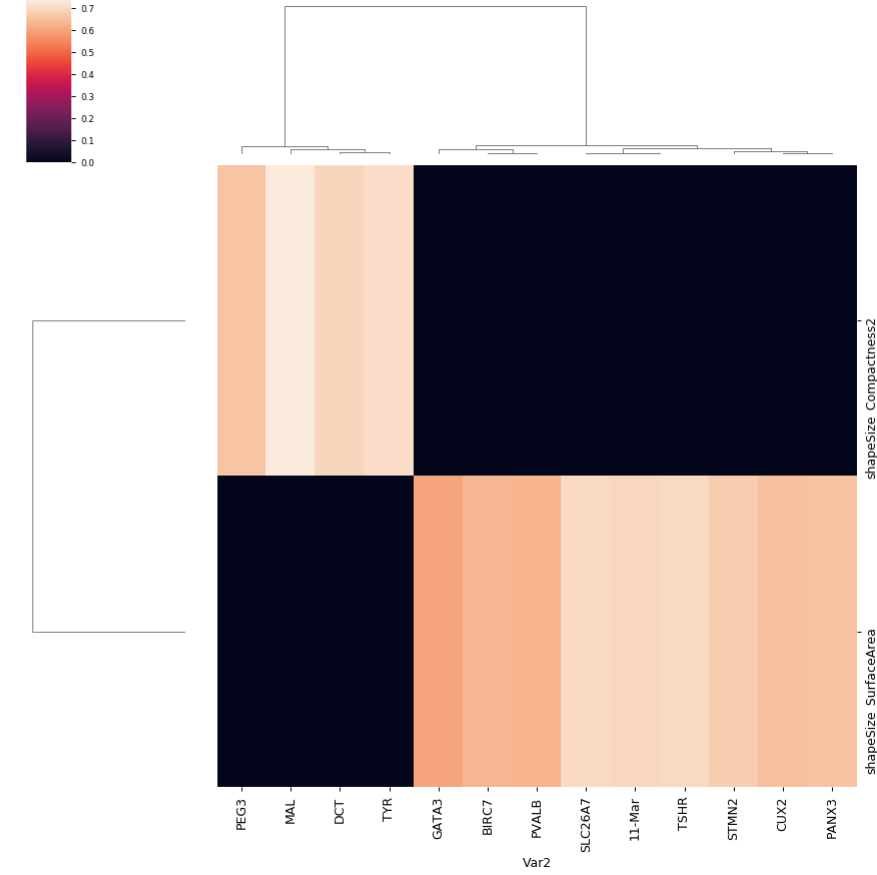
performing Stand_scaler normalization for imaging features

performing Stand_scaler for gene features

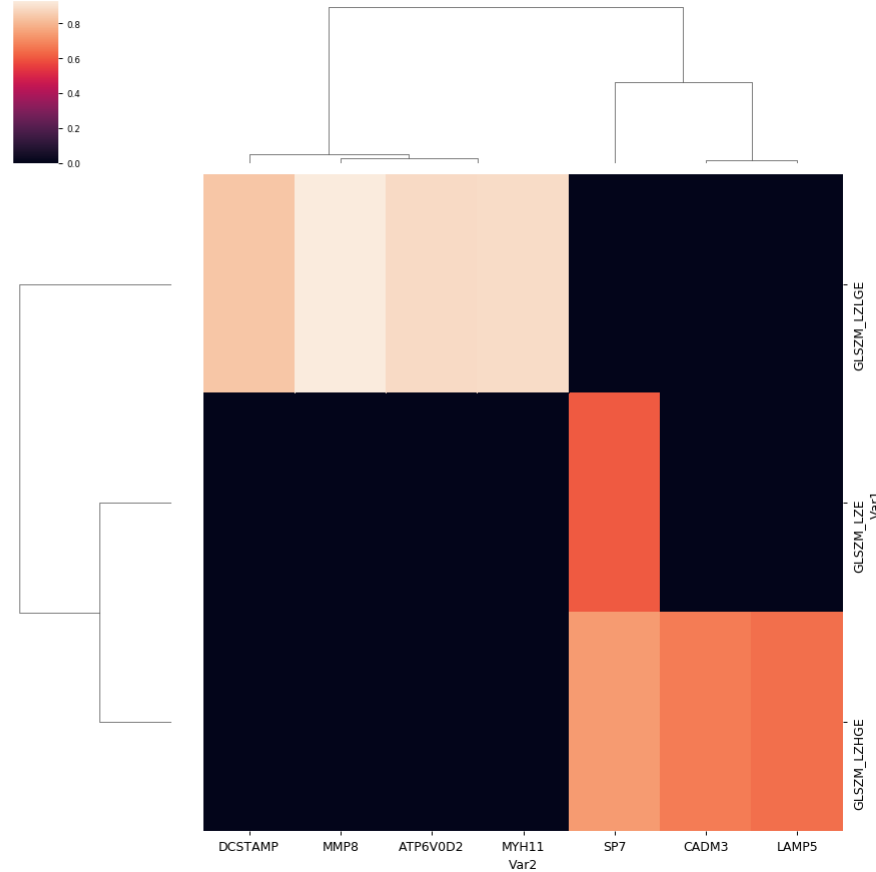
Top significant correlations (FDR_adjusted_pValue<0.05) for GLRLM features



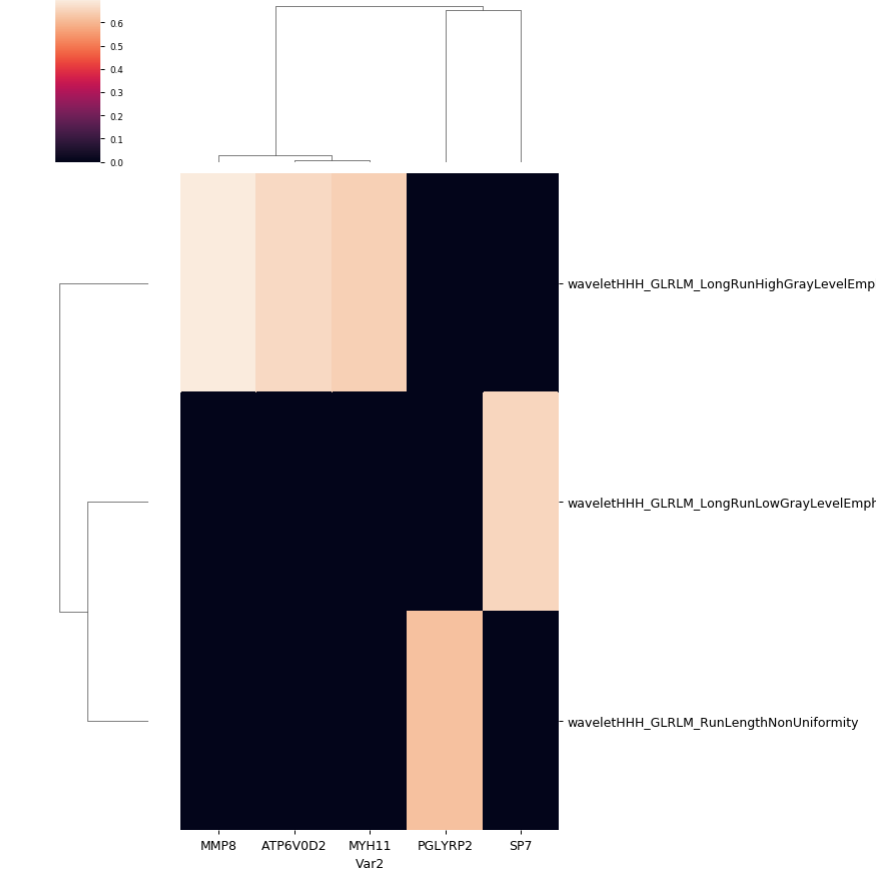
Top significant correlations (FDR_adjusted_pValue<0.05) for shapeSize features



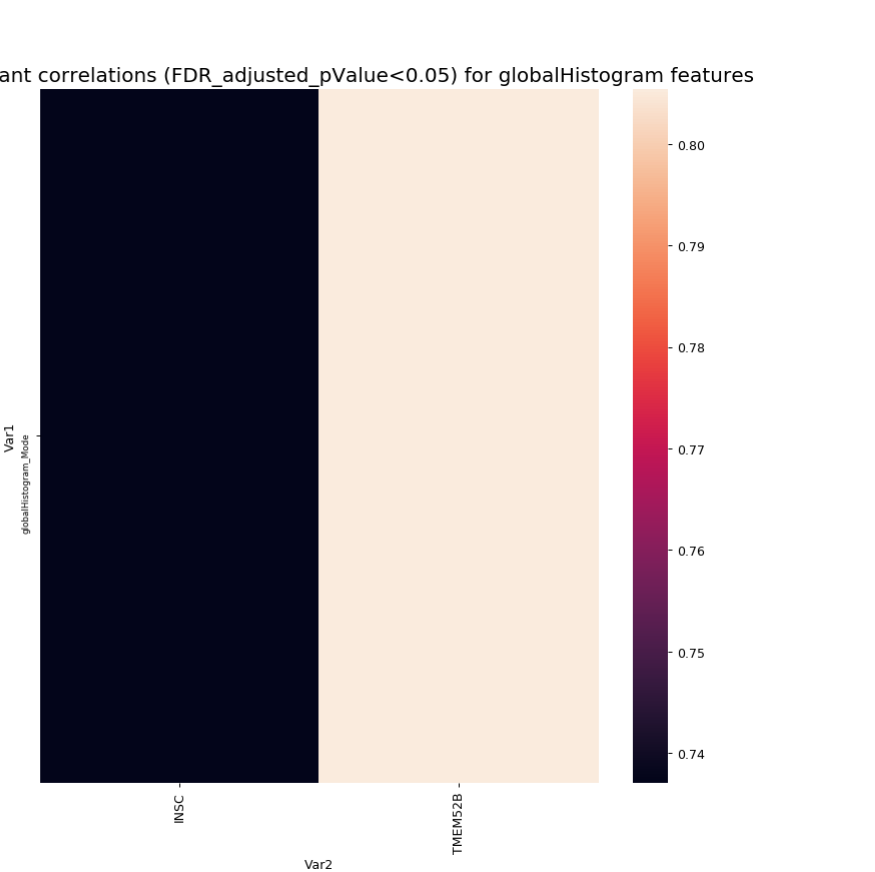
Top significant correlations (FDR_adjusted_pValue<0.05) for GLSZM features



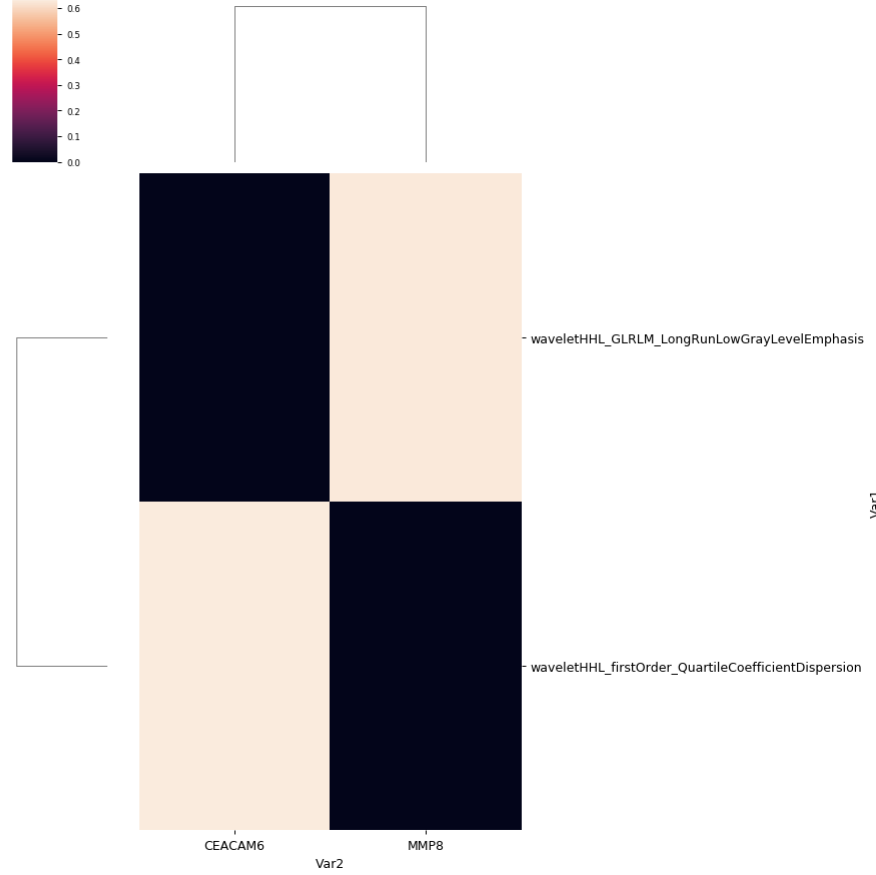
Top significant correlations (FDR_adjusted_pValue<0.05) for waveletHHH features



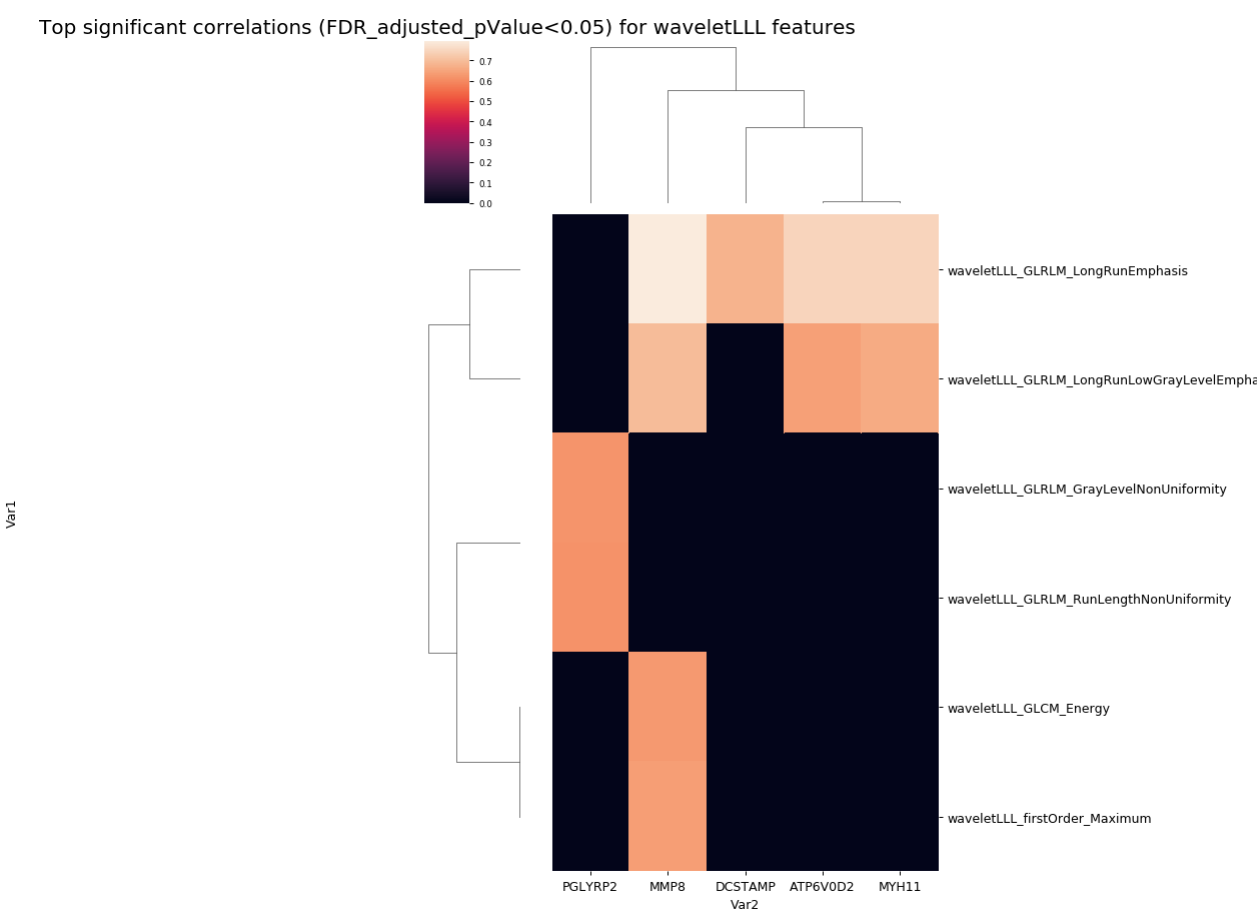
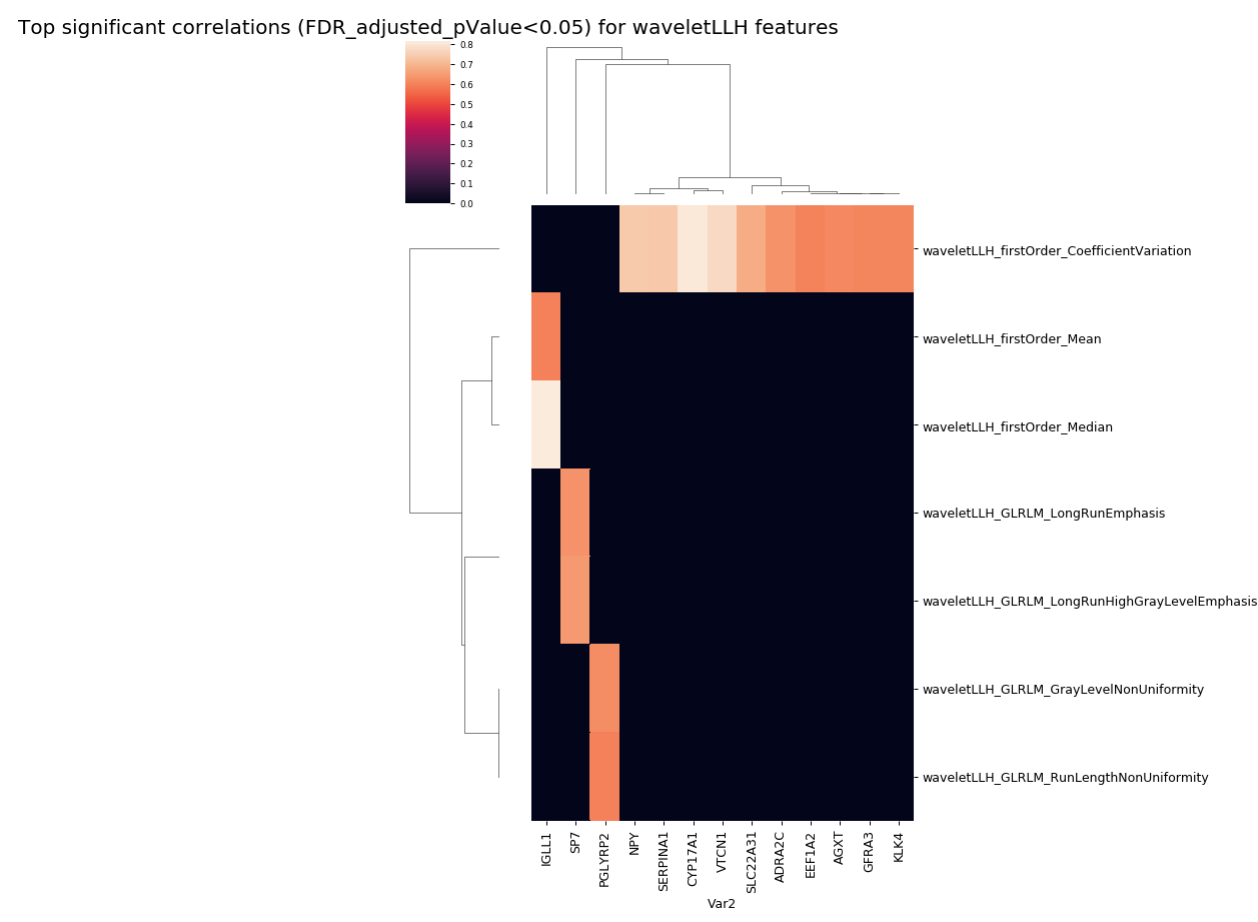
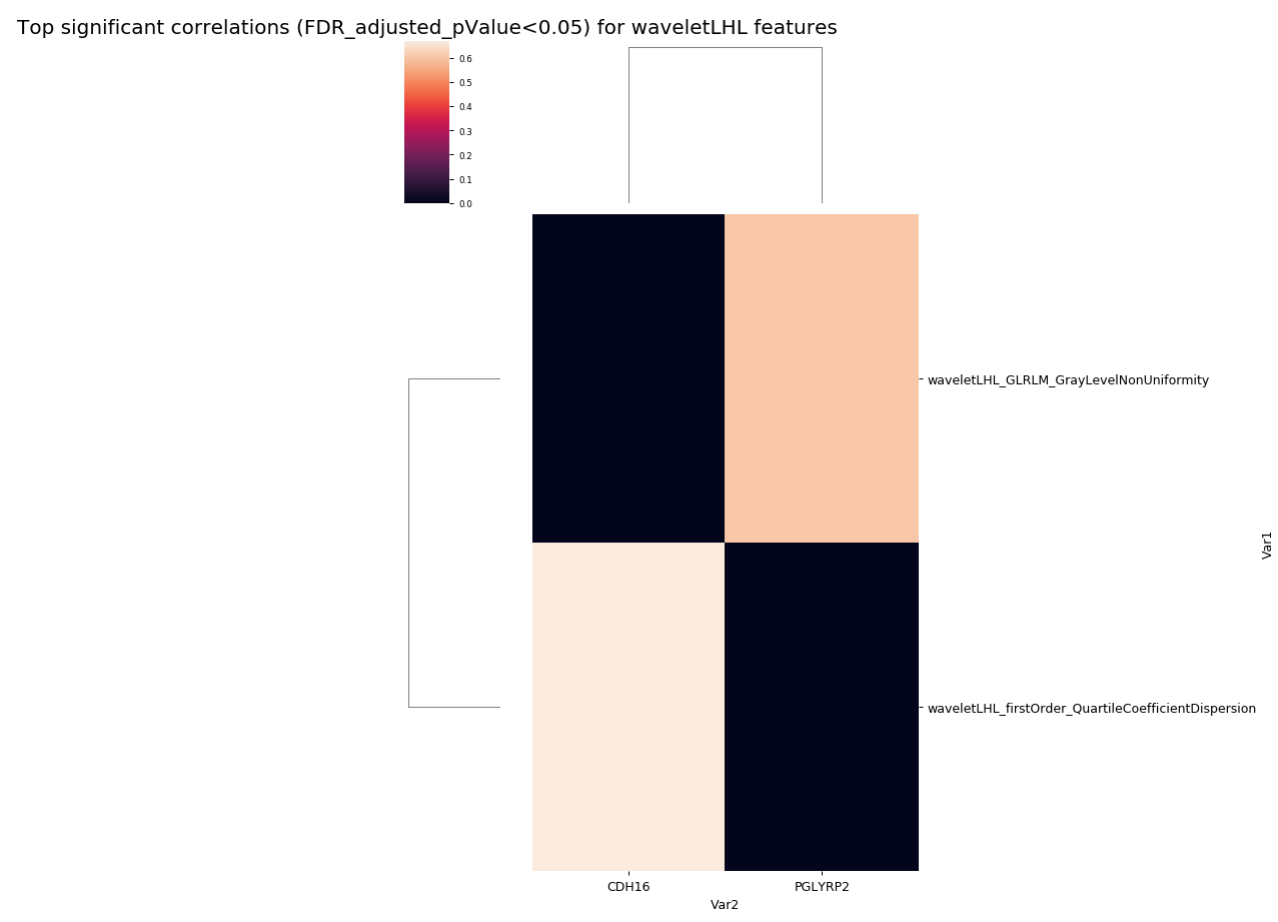
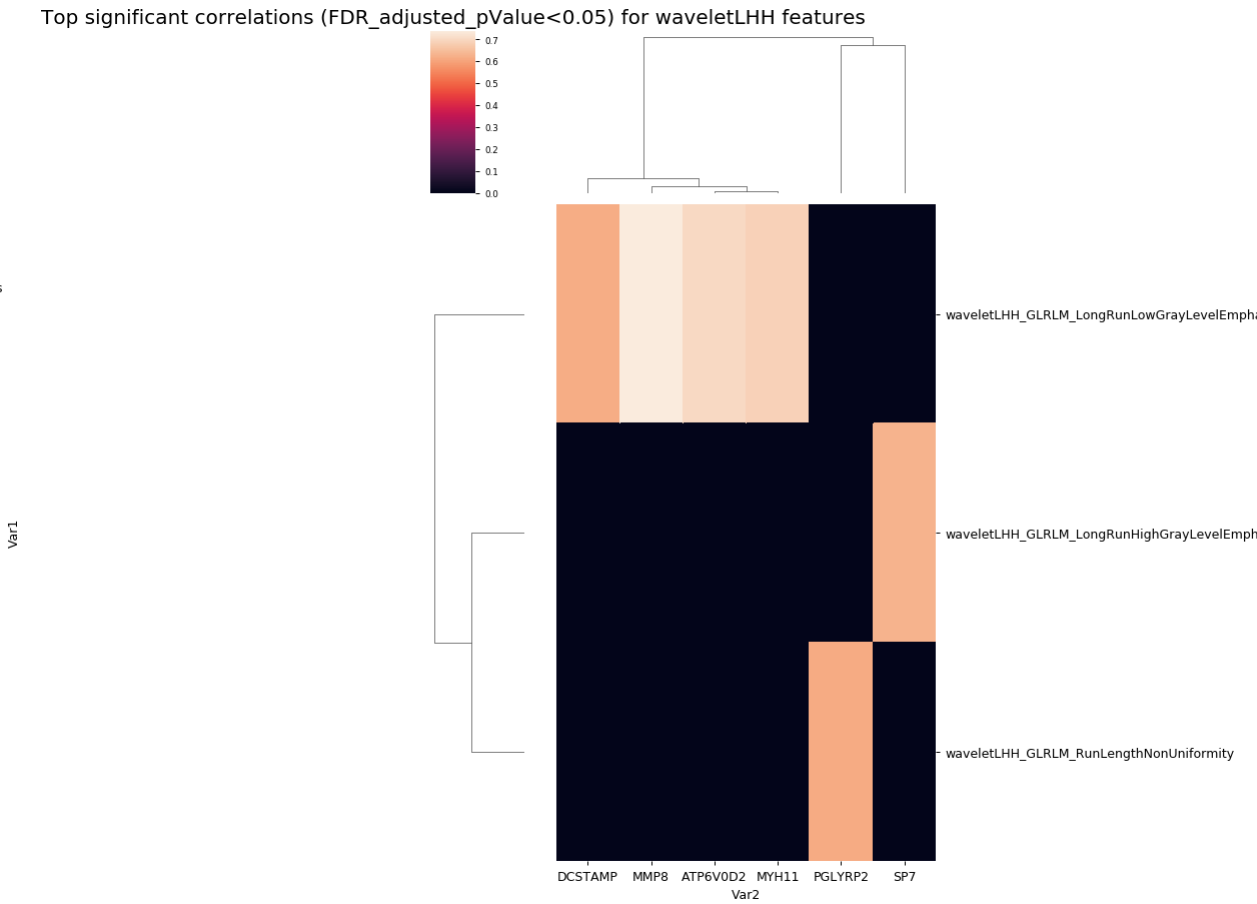
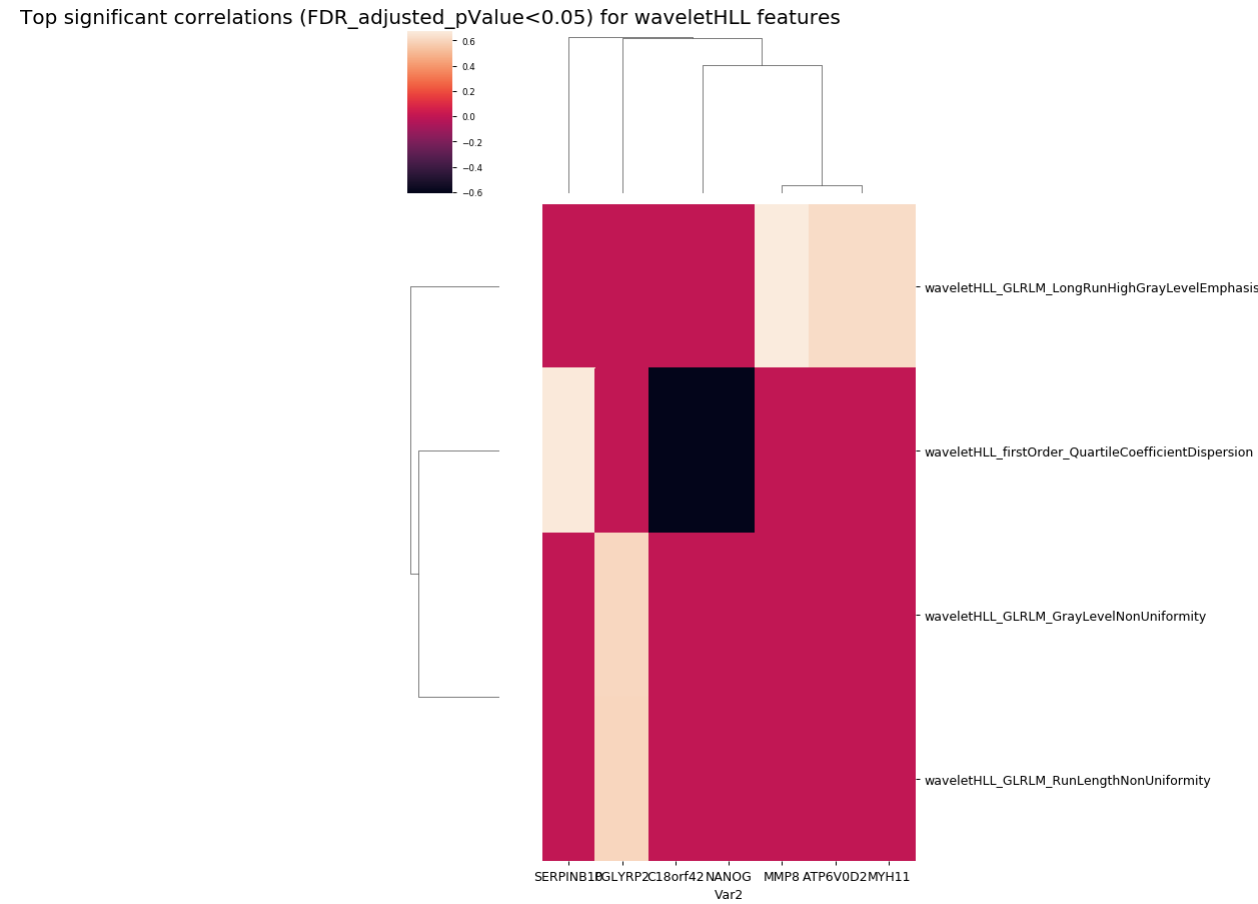
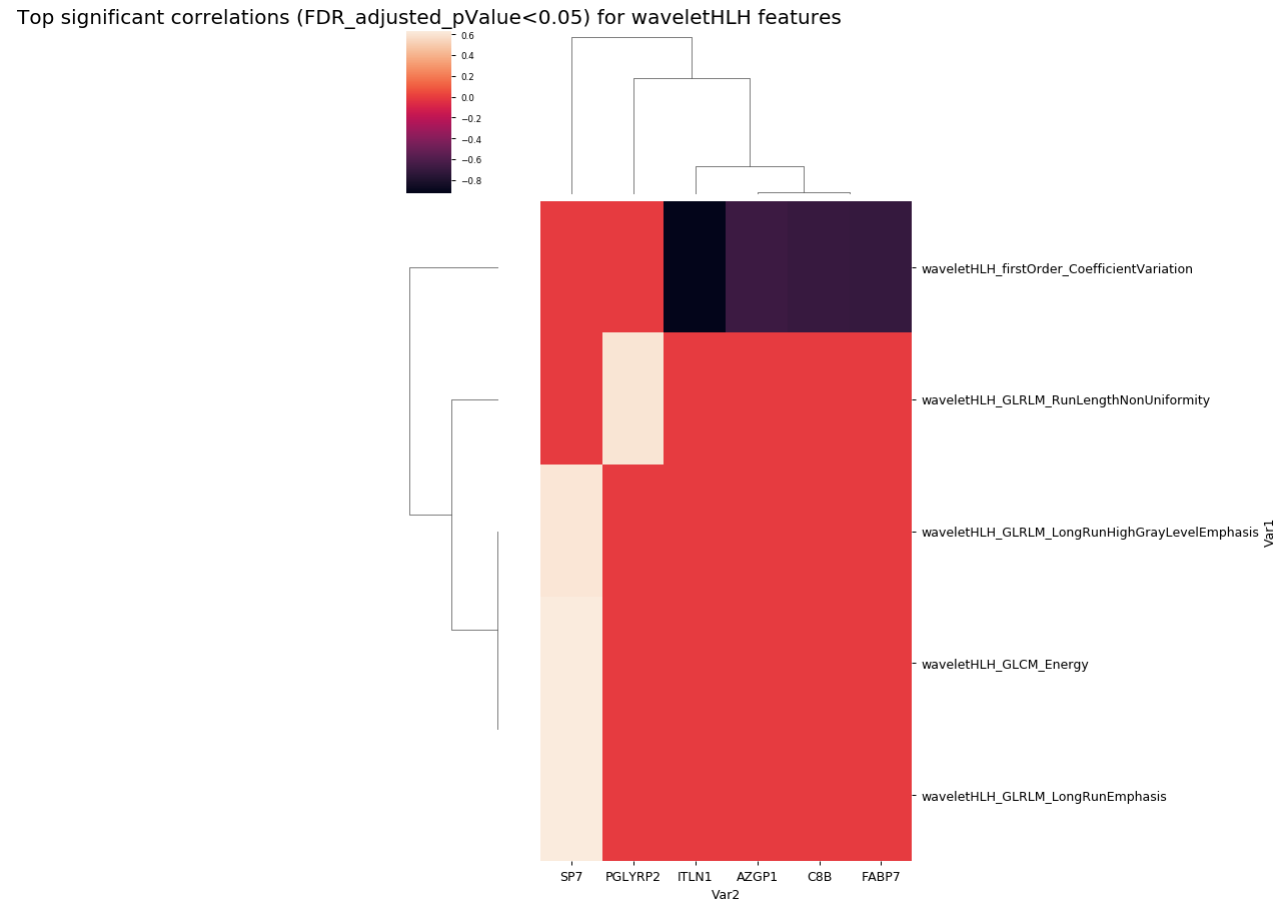
Top significant correlations (FDR_adjusted_pValue<0.05) for globalHistogram features



Top significant correlations (FDR_adjusted_pValue<0.05) for waveletHHL features



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



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

Below is the list of imaging features

['GLRM_LongRunEmphasis', 'GLRM_LongRunLowGrayLevelEmphasis', 'GLRM_RunLengthNonUniformity', 'GLSZM_LZE', 'GLSZM_LZHGE', 'GLSZM_LZLGE', 'globalHistogram_Mode', 'shapeSize_Compactness2', 'shapeSize_SurfaceArea', 'waveletHHH_GLRM_LongRunHighGrayLevelEmphasis', 'waveletHHH_GLRM_LongRunLowGrayLevelEmphasis', 'waveletHHH_GLRM_RunLengthNonUniformity', 'waveletHLL_GLRM_LongRunLowGrayLevelEmphasis', 'waveletHLL_GLRM_LongRunHighGrayLevelEmphasis', 'waveletHLL_GLRM_RunLengthNonUniformity', 'waveletLHH_GLRM_LongRunHighGrayLevelEmphasis', 'waveletLHH_GLRM_LongRunLowGrayLevelEmphasis', 'waveletLHH_GLRM_RunLengthNonUniformity', 'waveletLHL_GLRM_LongRunEmphasis', 'waveletLHL_GLRM_LongRunHighGrayLevelEmphasis', 'waveletLHL_GLRM_RunLengthNonUniformity', 'waveletLLL_GLRM_LongRunEmphasis', 'waveletLLL_GLRM_LongRunLowGrayLevelEmphasis', 'waveletLLL_GLRM_RunLengthNonUniformity', 'waveletLLL_GLRM_GrayLevelNonUniformity', 'waveletLLL_firstOrder_Maximum']

Below is the list of gene features

['IL13', 'ADRA2C', 'AGXT', 'ATP6V0D2', 'AZGP1', 'BIRC7', 'C18orf42', 'C1QL1', 'C8B', 'CADM3', 'CDH16', 'CEACAM6', 'CUX2', 'CYP17A1', 'DCSTAMP', 'DCT', 'DPPA3', 'EEF1A2', 'FABP7', 'G6PC', 'GATA3', 'GFR3', 'HISTH3G', 'HISTH4D', 'IGLL1', 'INSC', 'ITLN1', 'KLR4', 'LAMP5', 'LIN28A', 'MAL', 'MMP8', 'MMP9', 'MYH11', 'NANOG', 'NPY', 'PANK3', 'PEG3', 'PGLYRP2', 'POU5F1', 'PRDM14', 'PSMB1', 'PVALB', 'SERPINA1', 'SERPINB1', 'SLC22A31', '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 : DecisionTree

Cross Validation Metrics:

Parameters: cv=3 scoring=neg_mean_squared_error

Cross validation score:1.0885822222

Model Parameters:

presort:False

splitter:best

min_impurity_decrease:0.0

max_leaf_nodes:None

min_samples_leaf:1

min_samples_split:2

min_weight_fraction_leaf:0.0

criterion:mse

random_state:None

min_impurity_split:None

max_features:None

max_depth:None

Min Square Error for the Model

MSE of train_eval set:0.0

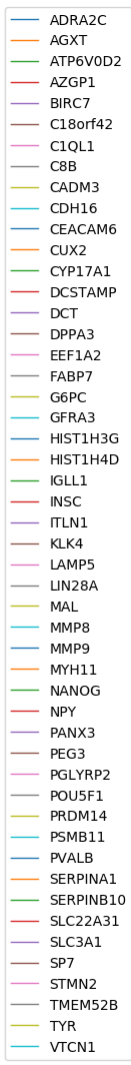
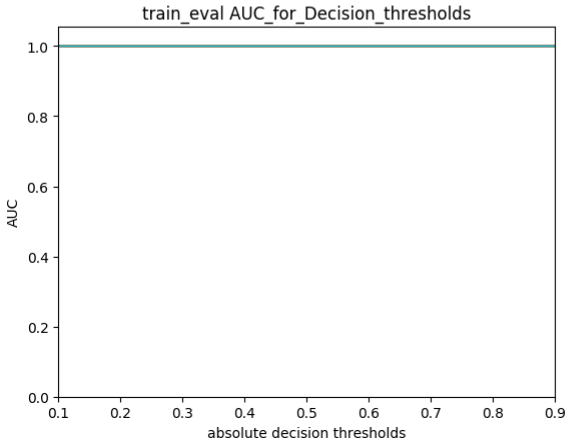
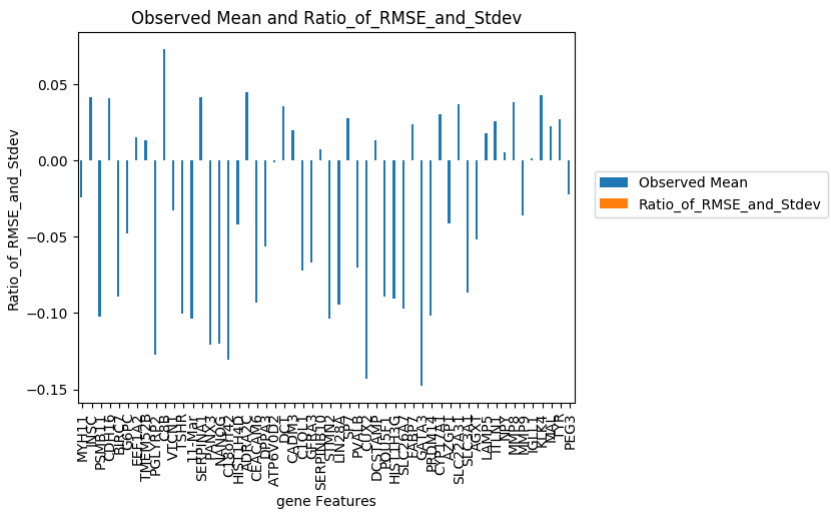
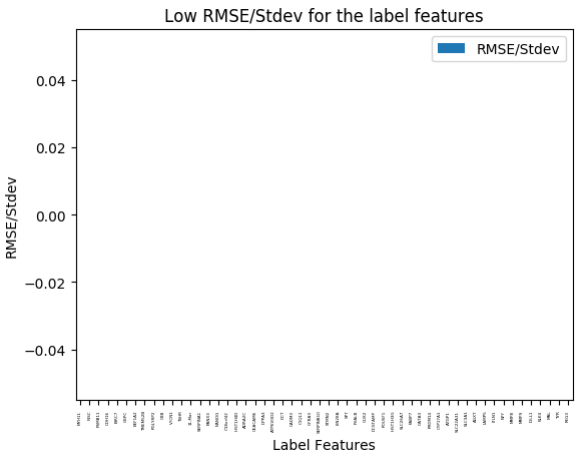
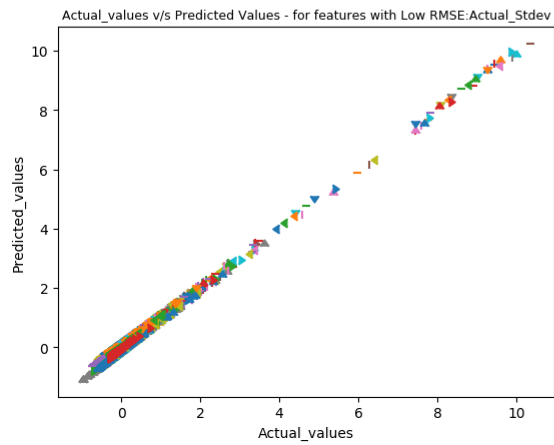
No. of features showing LOW 'RMSE/Stdev' (≤ 1.0): 54

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

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



Model evaluation for Test data

Min Square Error for the Model

MSE of test_eval set:2.09923242643

No. of features showing LOW 'RMSE/Stdev' (≤ 1.0): 5

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

No. of features showing HIGH 'RMSE/Stdev' (> 1.0): 49

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

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

