

Name	Records	Features	Description
Dataset1_scaled_Raw Features_Imbalanced .csv	617	52 + Class = 53	<ul style="list-style-type: none"> Classes are imbalanced. Features are scaled. Missing values- None Correlation >0.8 features discarded. Id column dropped. Needs to train/test splitted before proceeding
Balanced_Train_data _raw_features.csv	174 (87+87)	52 + Class = 53	<ul style="list-style-type: none"> Its a training dataset. Class distribution balanced. Features are scaled. Missing values- None Correlation >0.8 features discarded. Id column dropped.
Unbalanced_Test_data _raw_features.csv	124 0 103 1 21	52 + Class = 53	<ul style="list-style-type: none"> Corresponding Test Dataset for the above training dataset - Balanced_Train_data_raw_features.csv Unbalanced to see if the model identified the imbalanced nature of the dataset post training.

Name	Records	Features	Description
dataset_65_features_train.csv	174 (87+87)	65 + Class = 66	<ul style="list-style-type: none"> • Combined Train dataset with 35 PCA comp and 30 original features • Class distribution balanced. • Features are scaled. • Missing values- None • Correlation >0.8 features discarded. • Id column dropped.
dataset_65_features_test.csv	124 0 103 1 21	65 + Class = 66	<ul style="list-style-type: none"> • Corresponding Test Dataset for the above training dataset - dataset_65_features_train.csv
sparse_matrix_train.csv	174 (87+87)	90 + Class = 91	<ul style="list-style-type: none"> • Combined Train dataset with 50 Truncated SVD comp and 40 original features • Class distribution balanced. • Features are scaled. • Missing values- None • Correlation >0.8 features discarded. • Id column dropped
sparse_matrix_test.csv	124 0 103 1 21	90 + Class = 91	<ul style="list-style-type: none"> • Corresponding Test Dataset for the above training dataset - sparse_matrix_train.csv