



Comparison of classifiers and pre-processing methods (Gaucher dataset). Each combination of pre-processing method and peak selection was ranked by its average classification accuracy on 500 test sets (size of training sets: 27, size of test sets: 12) for each classifier. The heatmap gives a colour coding of the ranks from 1 (highest accuracy, red) to 6 (lowest accuracy, light yellow). Columns of the heatmap are ranked by their average rank over all classifiers, with Ciphergen pre-processing using setting A getting the highest rank. Classifiers are ordered by their average rank over all pre-processing combinations, with SVM being the best ranked classifier.