GlyStruct: glycation prediction using

structural properties of amino acid residues

Summary

The authors of the study introduced a new machine learning method called GlyStruct and Support Vector Machine (SVM) as a classifier. Using the classifier, they predict the glycated and non-glycated lysine residues using properties of amino acid residues. Where glycation is a Post-translational Modifications (PTM) that has a significant clinical relevance attributed to a correlation with increase blood sugar and metabolic morbidity detection. They collected the dataset from online databanks that is widely used CPLM database. Dataset formulated to make it suitable for training classifiers and an appropriate cross-validation scheme for prediction. K-nearest neighbor (KNN) filter is used to resolve the imbalance in dataset. It contains 235 Glycated and 303 non-Glycated lysine residues. For feature extraction, they considered three attributes are the secondary structure, local backbone torsion angles and accessible surface area (ASA). They used`1 k-fold cross validation like 6-, 8- and 10-folds for evaluation of Glystruct predictor. GlyStruct provides approximately 10% better performance in compare to benchmark method of Gly-PseAAC. For 10-fold cross validation, the performance of GlyStruct on the metrics, sensitivity, specificity, accuracy and Mathew’s correlation coefficient (MCC) were 0.7013, 0.7989,0.7562 and 0.5065.