Banlang Niyomsak 2010: Sequential Associative Classification for Enzyme Subfamily Prediction. Master of Engineering (Computer Engineering), Major Field: Computer Engineering, Department of Computer Engineering. Thesis Advisor: Associate Professor Kitsana Waiyamai, Ph.D. 69 pages.

Enzyme subfamily prediction is nowadays a very active bioinformatics research topic. Using statistics to build predictor, enzyme subfamily prediction methods are very accurate. However those methods generate non-explainable predictor models and order between features is not taken into account for prediction model construction.

In this research work, we propose a Sequential Associative Classification (SAC) for enzyme subfamily prediction. SAC integrates sequential pattern mining with associative classification to perform prediction task. Background knowledge about biding and active site motifs is integrated in the sequential pattern mining step. Associative classification is extended by taking into account the order between features during the prediction model construction step. Experimental results using SWISS-PROT sequence dataset show that SAC is able to generate a very easy to understand descriptive predictor in the forms of class sequential pattern rules, and a very high accuracy of 70.95%.

		/ /
Student's signature	Thesis Advisor's signature	