1154-62-1616 **Xu (Sunny) Wang*** (xwang@wlu.ca), Department of Mathematics, 75 University Avenue West, Waterloo, Ontario N2L 3C5, Canada, and **Hugh Chipman** (hugh.chipman@acadiau.ca), Department of Mathematics and Statistics, Acadia University, 15 University Ave, Wolfville, NovaScotia B4P 2R6, Canada. *A New Constrained Mixture Models for Drug Discovery Data with Innovative Iterative Algorithms.*

It has been well-known that Statistics has played a very important role in drug discovery: facilitate and speed up the drug finding process. Statistical learning in drug discovery seeks a good classifier that separates chemical compounds into active and inactive classes. The active compounds are those that inhibit disease virus and are taken as drug candidates. However, the characteristics of drug data imply many challenges for structure modeling and identification of active compounds. Among these characteristics are the rarity of active compounds, the large volume of compounds tested by high-throughput screening, sub-set governed activity and the complexity of molecular structure and its relationship to activity. Due to the challenges of drug discovery data, we develop a new statistical learning model based on mixture models: Constrained Mixture Discriminant Analysis (CMDA) model. This method is designed to catch multiple mechanisms that lead to activity, explore the subsets of descriptors and be easily interpreted (e.g. identify important descriptors). The Expectation-Maximization, is used to estimate the parameters of the CMDA model. Comparing to the popular MclusteDA model, CMDA outperforms MclustDA in terms of the parameter estimation and the active compound identification. (Received September 16, 2019)