1056-90-2129   **Ye (Alice) Tian*** (`ytian@math.wsu.edu`), 1445NE, North Fairway, Dr L45, pullman, WA 99163.
Optimization in Developing Prediction Scoring Function for Mutagenesis Solubility.

Mutagenesis is the process of replacing one or more amino acids in a protein by alternative amino acids to generate a mutant protein. Computational techniques can be invaluable to predicting the changes to the properties without having to actually make the mutations.

We use concepts from computational geometry to develop a weighted scoring function to predict the change of protein solubility due to mutations. The scoring function captures both protein sequence and structure information. We present two Linear Programming (LP) models to train the scoring function. The first one is a two-phase model that tries to distinguish boundary between increased solubility and decreased solubility instances. The second model is trying to minimize the deviations of all instances from hitting the target values. The second model also shows robustness for different target values. Utilizing Leave-One-Out Cross Validation method for training and prediction, a prediction accuracy 94.5% is achieved.

Compared to standard learning methods such as Support Vector Machines (SVM), the LP method offers more flexibility, and is less sensitive to outliers as the objective function is linear. (Received September 23, 2009)