

Title of Research:

12_PT01-06

Sophisticated hazard prediction by active QSAR modeling

Principal Investigator:

Yoshimasa Takahashi (Professor, Dept. Comp. Sci. Eng., Toyohashi Univ. Tech.)

Collaborators:

Hiroaki kato (Asoc. Professor, Dept. Comp. Sci. Eng., Toyohashi Univ. Tech.)

Tomoya Yamazaki (Master student, Toyohashi Univ. Tech.)

Keisuke Saito (Master student, Toyohashi Univ. Tech.)

Summary of Research:

In preceding work, we proposed a technique for active QSAR Modeling that is based on active sampling of a temporary training data set, which gives us higher performance in the prediction. However, the prediction result depends on the number of compounds for the temporary training set used for making the model. And, it is difficult to know the optimal number of the neighbors to be used, in advance. In the present work, we employed a threshold of the similarity at exploring the neighbors but not the number of neighbors to be searched. Computer experiment showed us that the method with the threshold of similarity gives a better performance. Besides, it was shown that a QSAR model obtained from the whole data could give us better prediction when the appropriate neighbors are not available enough. Alternatively the performance of the method also depends on the local data structure around a query of interest. But, it is impossible to see a data structure of the local space around the query in advance. We employed kNN, RMSST (Rooted Minimum Spanning Sub-Tree) and Centroid method to explore the neighbors. Once, we generated the QSAR models with the different training sets that are obtained by the different neighbor searching methods, and we evaluated the statistical performance of the models. We used the best approximation model among them. Computational experiment with a data set of toxic chemicals suggested that the current approach can provide us much better predictions for the case.

Timeline: November 1, 2012 -

Topics:

The Second New LRI Workshop, Tokyo, Aug., 2013

Publications:

- 1) Tomoya Yamazaki and Yoshimasa Takahashi, Toxicity prediction of chemical substances by Active Sampling :Introduction of cutoff value by similarity, Proc. of the 22nd Symposium on Environmental Chemistry, 2013, July, Tokyo.
- 2) Tomoya Yamazaki and Yoshimasa Takahashi, Toxicity prediction of chemical substances by Active Sampling: Neighbor search method and model selection, Proc. of the 41th Symposium on Structure-Activity Relationships, 2013, November, Osaka.