Open Source data mining in chemoinformatics

Egon Willighagen

Published November 2, 2005

Citation

Willighagen, E. (2005). Open Source data mining in chemoinformatics. In *chem-bla-ics*. chem-bla-ics. https://doi.org/10.59350/gc4hw-5k265

Keywords

Iccs, Cheminf

Abstract

On the 7th International Conference on Chemical Structures Jeroen Kazius has a poster on finding discriminative substructures, that is, molecular fragments which can be discriminate between two acitivity classes. The software is released as Gaston , is written in C++ and has the GPL license.

chem-bla-ics

On the 7th International Conference on Chemical Structures Jeroen Kazius has a poster on finding discriminative substructures, that is, molecular fragments which can be discriminate between two acitivity classes. The software is released as Gaston , is written in C++ and has the GPL license.

Later I encountered MoSS which has the same goal, but uses a different algorithm. MoSS is written in Java and uses the LGPL license. MoSS reads STN and SMILES as input, which might not be optimal for all users, so a CDK port comes to mind.