Matt gave this presentation at the ACS National Spring Meeting 2012.

Abstract

Automatic QSAR model building methods are now readily available and have been successfully applied to a range of compound properties (solubility, logP, Blood-brain barrier penetration etc.). However, good ligand-based models of target potency are less common and have traditionally been dependent on the application of 3D and structure-based approaches. However, the availability of good public-domain data sources and the latest machine learning techniques in an automated framework have enabled us to carry out a comprehensive study of the potential for building 2D ligand-based models of target potency. We will discuss the results of automatically applying multiple QSAR model building techniques (PLS, Radial Basis Functions, Gaussian Processes, Random Forests) to over 70 data sets across a range of target classes using data from the ChEMBL database. We will explore the effects of the quality and quantity of data, modelling method and model domain of applicability on the accuracy of prediction.

These are the slides that Matt presented.

A copy of Matt's slides is available as a PDF file.