A Correlation-Based Approach to Attribute Selection in Chemical Graph Mining

Takashi OKADA

Data mining often encounters a problem with a huge number of descriptive features. Authors have analyzed structure activity data of dopamine antagonists, where we have to select useful features out of numerous fragments extracted from chemical structures. Correlation coefficients among categorical attributes are used to select attributes. Rules obtained by the cascade model were evaluated from chemists' point of view, and the importance of attribute selection was confirmed.