Machine learning model for activation free energy and TS geometry of pericyclic reactions

Yujiro Matsuo
TS Technology

Recently, there has been a focus on predicting transition state (TS) structures and activation free energies derived from quantum chemical calculations using machine learning. Once a predictive model is established, machine learning enables much faster numerical predictions compared to quantum chemical methods. Using this approach, it becomes possible to rapidly rank synthetic routes generated by synthetic route design systems (SRDS), such as AIPHOS/KOSP/TOSP, based on their quantum chemical priorities. This would significantly shorten the development time for functional chemicals. In this study, we developed a machine learning model to predict the DFT-calculated values of ΔG‡ and the atomic distances at the reaction center of the TS structure, trained on a large dataset of quantum chemical calculation results. We extracted 347 TS data from the Diels-Alder reaction in our transition state database (TSDB), which compiles quantum chemical data, and used these as training data. The predictive model, built using a Counter-Propagation Neural Network (CPNN), an extension of Kohonen's Self-Organizing Map, achieved an RMSE of 2.304 kcal/mol for ΔG‡ and an RMSE of around 0.1 Å for the atomic distances at the reaction center of the TS structure, providing sufficient accuracy for pathway screening in both cases.