Prediction of Excitation States of Cerium Complexes via Machine Learning

Kariya Kosuke, Keio University

Recently, cerium (Ce) complexes with 4f-5d transition properties have been getting attentions as photoreduction catalysts. However, their ligand designs are difficult because the relationship between the optical functions and the ligands is unknown. In addition, in silico screening is challenging due to the large computational cost. To overcome these problems, we focused on machine learning (ML) models that could predict the optical properties from computationally inexpensive parameters such as molecular fingerprints and ground state properties. To build ML models, 1427 Ce complexes were gathered from the CCDC database. We constructed two ML models to precisely predict the excited state of Ce complexes. The first is classification models to determine whether Ce complexes have the 4f-5d transitions properties or not, and the second is regression models to predict the excitation energy of the 4f-5d transition. Finally, we extracted the knowledge for designing new Ce complexes with desired excited-state characters and high excitation energies by analyzing the contribution of each explanatory variable like properties or substructures to the constructed ML models.