CH 434 Machine Learning in Chemistry

Introduction to encode molecules in a Format Readable and Processable by Computers: Cheminformatics learning, smiles representation, RDKIT, molecular to vector representation; Introduction to regression and classification style statistical analysis: types of regression and classification, data visualizations, simple chemical properties selection, feature identification, descriptor findings; Data-driven modeling: data and standardization, training set vs testing set, model validations, evaluating the model accuracy, simple chemical properties selection, feature identification, descriptor findings; Introduction to Deep Learning, Generative AI, Graph neural network; Property prediction models: example code for libraries (Pandas, Numpy, Matplotlib etc.), example code for materials, catalysis, potential energy surface, and reaction prediction.

Text

1. Hugh M Cartwright, Machine learning in Chemistry: The Impact of artificial Intelligence, 1st Edition, Royal Society of Chemistry.2020
2. Jahan B. Ghasemi, Machine Learning and Pattern Recognition Methods in Chemistry from Multivariate and Data Driven Modeling,1st Edition, Elsevier, 2022

References

1. Etienne Bernard, Introduction to Machine Learning, 1st Edition, Wolfram Media Inc, 2021
2. Mark Lutz, Learning Python: Powerful Object-Oriented Programming, 5th Edition, O'Reilly Media, Inc, 2013
3. Aurélien Géron, Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow, 2nd Edition, O'Reilly Media, Inc, 2019