Predicting Molecular Properties

Sunghwan Choi

National Institute of Supercomputing and Network
Korea Institute of Science and Technology Information
Daejeon, 34141, Republic of Korea

In this presentation, I will share some experience to participate machine learning competition whose goal is to accurately predict spin coupling constants of the given quantum chemical results.[1] A spin coupling constant is an important factor to determine structural and dynamical properties of molecule from NMR spectrum but the accurate prediction needs a large amount of time compared to an inference of machine learning model. 2749 teams from all around world joined this competition and my team got 3rd prize.

The 3rd solution will be also addressed in this talk. Even though the largest molecule in the training set is a nonane(C_9H_{20}) molecule but a conventional graph-based algorithms are not successfully trained due to the huge size of training set. In order to effectively reduce computational cost with preserving the accuracy, my team applied the attention only architecture. The details of model and some practical tips to enhance performance including seed ensemble and pseudo-labeling techniques will be discussed.

[1] Homepage of Predicting Molecular Properties https://www.kaggle.com/c/champs-scalar-coupling/overview (2020), last accessed 16. Jan. 20