Intro to Feature Representation in Virtual Screening

Shengchao Liu, Gitter Group
Feature Representation

1. Raw Molecule Representation (Graph CNN)
   a. atom info
   b. bond info

2. SMILES (RNN, CNN, RNN+CNN)
   a. string, a sequence of characters
   b. lack of structural info

3. Morgan Fingerprints/ECFP (Dense NN, classical ML)
   a. hashing will lose more information
Graph CNN Framework
Related Work

Related Work

- Vanilla LSTM, (Keck Paper)

More oftenly used in molecule generation tasks, like GAN and AE.