

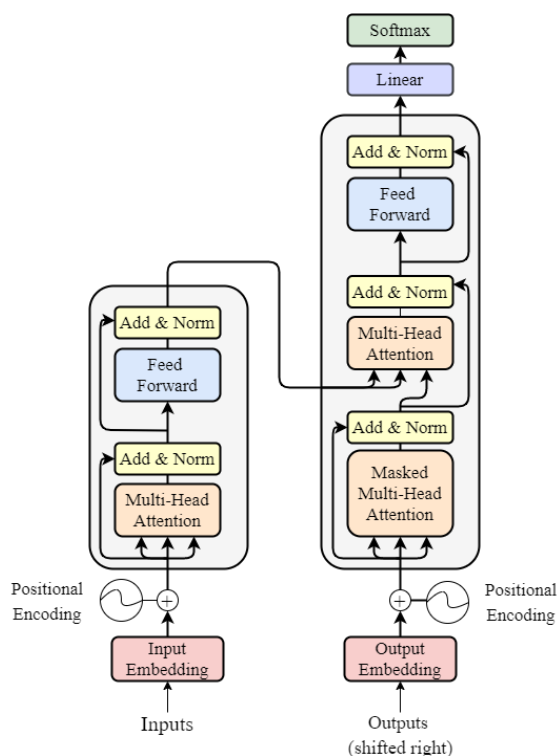
Transformatorių neuroninių tinklų taikymas DFT molekulių struktūrų optimizavimo skaičiavimams spartinti

Using transformer neural networks to speed up molecular structure optimization at DFT level

Rokas Garbačas, Stepas Toliautas

Vilniaus Universitetas, Fizikos fakultetas, Cheminės fizikos institutas, Saulėtekio al. 9, III rūmai, LT-10222, Vilnius
rokas.garbacauskas@ff.stud.vu.lt

Transformatorių neuroniniai tinklai (1 pav.) yra santykinai nauja inovacija mašininio mokymo srityje [1], kuri leido sukurti pastaruoju metu labai išpopuliarėjusius didelius kalbos modelius (ChatGPT, LLaMA ir kt.) bei pasiekti puikius rezultatus vertimo ir vaizdų generavimo sferose. Ta pati architektūra su tam tikromis modifikacijomis leido pasiekti stulbinančius rezultatus prognozuojant baltymų struktūrą iš amino rūgščių sekos [2]. Šių tinklų pagrindinė stiprybė yra duomenų užkodavimas juos aprašančiais vektoriais bei dėmesio mechanizmas, kuris leidžia modeliui atpažinti net labai subtilias koreliacijas tarp duomenų.



1 pav. Transformatorių neuroninio tinklo architektūra, pasiūlyta [1] straipsnyje.

Šio tipo tinklai puikiai tinka modeliuoti vienmates, t. y. išrikuojamas, sistemas. Efektyvus jų taikymas duomenims daugiamatėje erdvėje yra aktyviai nagrinėjamas klausimas. Molekulių struktūra dažniausiai optimizuojama trimatėje erdvėje, tad efektyvus šios architektūros pritaikymas molekulių struktūrai modeliuoti yra netrivialus uždavinys. Taip pat atviras klausimas yra ir cheminės informacijos

kodavimas vektoriniuose deskriptoriuose bei dėmesio mechanizme.

Įvairūs neuroniniai tinklai jau naudojami kvantinės chemijos skaičiavimuose ir leidžia atlikti molekulių struktūrų optimizaciją, skaičiuoti jėgas bei dipolinius momentus ir kt. su vartotojų lygio įranga žymiai sparčiau nei atliekant DFT skaičiavimus superkompiuterių mazguose [3], [4], [5]. Pagrindiniai neuronų tinklais grįstų metodų trūkumai lyginant su DFT skaičiavimais yra kiek prastesnės gautos energijos vertės, „juodosios dėžės“ veikimo pobūdis bei santykinai prastas duomenų generavimas cheminiams junginiams, nepatenkantiems į mokymo rinkinį.

Su transformatorių neuroninių tinklų architektūra mes norime paspartinti DFT skaičiavimus ir įvertinti, ar pavyksta gauti DFT lygio energijos/ teorinių spektrų tikslumą, skaičiavimams sunaudojant tik dalį įprastai reikalingų resursų. Verta paminėti, jog plečiantis atviros prieigos duomenų bazėse esančių kvantinės chemijos duomenų kiekiui bei įvairovei, automatiškai mažėja ir apmokytam modeliui neatpažįstamų junginių kiekis.

Reikšminiai žodžiai: DFT, molekulių struktūros optimizacija, mašininis mokymas, transformatorių neuroniniai tinklai

Literatūra

- [1] A. Vaswani et al., *Attention is All you Need*, Advances in Neural Information Processing Systems (2017).
- [2] J. Jumper et al., *Highly accurate protein structure prediction with AlphaFold*, Nature (2021).
- [3] O. T. Unke and M. Meuwly., *PhysNet: A Neural Network for Predicting Energies, Forces, Dipole Moments, and Partial Charges*, Journal of Chemical Physics and Computation (2019).
- [4] J. S. Smith, O. Isayev, and A. E. Roitberg. *ANI-1: an extensible neural network potential with DFT accuracy at force field computational cost*, Chemical Science (2017).
- [5] K. Yao, J. E. Herr, D. W. Toth, R. Mckintyre, and J. Parkhill. *The TensorMol-0.1 model chemistry: a neural network augmented with long-range physics*, Chemical Science (2018).