

# Transformatorių neuroninių tinklų taikymas DFT molekulių struktūrų optimizavimo skaičiavimams spartinti

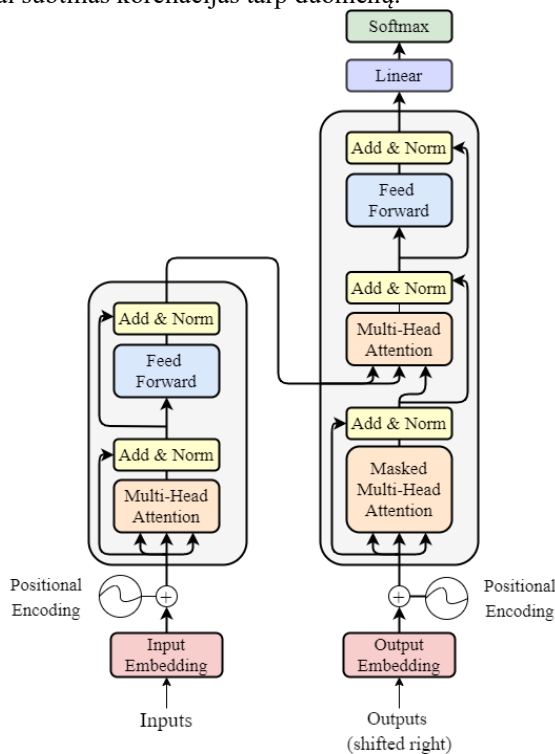
## Using transformer neural networks to speed up molecular structure optimization at DFT level

Rokas Garbačas, Stepas Toliautas

Vilniaus Universitetas, Fizikos fakultetas, Cheminės fizikos institutas, Universiteto g. 3, LT-01513 Vilnius

[rokas.garbacias@ff.stud.vu.lt](mailto:rokas.garbacias@ff.stud.vu.lt)

Transformatorių neuroniniai tinklai (1 pav.) yra santykinai nauja inovacija mašinio mokymo srity [1], kuri leido sukurti pastaruoju metu labai išpopuliarėjusius didelius kalbos modelius (angl. Large Language Models - LLMs) tokius kaip ChatGPT, LLaMA ir t. t. bei pasiekti puikius rezultatus vertimo ir vaizdų generavimo sferose. Ta pati architektūra su tam tikromis modifikacijomis Google padaliniui DeepMind leido pasiekti stulbinančius rezultatus prognozuojant baltymų struktūrą iš amino rūgščių sekos. Šių tinklų pagrindinė stiprybė yra duomenų užkodavimas juos aprašančiais vektoriais bei dėmesio mechanizmas, kuris leidžia modeliui išmokti net labai subtilias koreliacijas tarp duomenų.



1 pav. Transformatorių neuroninio tinklo architektūra pasiūlyta [1] straipsny.

Šie tinklai puikiai tinka modeliuoti viendimenses, t. y. išrikiuojamas, sistemas. Efektyvus jų taikymas duomenims daugiamačėje erdvėje yra atviras 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 skirtingi neuroniniai tinklai jau naudojami

kvantinės chemijos skaičiavimuose ir leidžia atlikti molekulių struktūrų optimizaciją, skaičiuoti jėgas bei dipolinius momentus ir kt. su vartotojų lygio aparatu žymiai sparčiau nei DFT skaičiavimai užtrunka superkompiuterių mazguose [2], [3], [4]. Pagrindiniai šių metodų trūkumai lyginant su DFT skaičiavimais yra kiek prastesnės gautos energijos vertės, juodos dėžės efektas bei palyginus prastas duomenų generavimas išėjus už treniravimo procese buvusių cheminių junginių rato.

Su transformatorių neuroninių tinklų architektūra mes norime pagreitinti DFT skaičiavimus ir gauti DFT lygio energijos tikslumą skaičiavimams sunaudojant mažą dalį pilniems DFT skaičiavimams reikalingų skaičiavimo/laiko resursų. Plečiantis atviros prieigos duombazėse esančių kvantinės chemijos duomenų kiekiui bei įvairovei automatiškai mažėja ir trečiosios iš problemų dydis.

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

### Literatūra

- [1] Vaswani et al., *Attention is All you Need*, Advances in Neural Information Processing Systems (2017).
- [2] 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).
- [3] 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*, 8(4):3192–3203, 2017
- [4] 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*, 9(8):2261–2269, 2018.