VIDY Reading Group
Drug Discovery

Longfeng Wu

CS, Virginia Tech

Jan 10th, 2024

VLOG

COMPUTER SCIENCE
VIRGINIA TECH

GEOLDM

Algorithm 2 Sampling Algorithm of GEOLDM

Input: decoder network $\mathcal{D}_\xi$, denoising network ϵ_θ
$\mathbf{z}_{\mathrm{x},T}, \mathbf{z}_{\mathrm{h},T} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$
for t in $T, T-1, \cdots, 1$ **do**
  $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$ {Latent Denoising Loop}
  Subtract center of gravity from $\boldsymbol{\epsilon}_\mathrm{x}$ in $\boldsymbol{\epsilon} = [\boldsymbol{\epsilon}_\mathrm{x}, \boldsymbol{\epsilon}_\mathrm{h}]$
  $\mathbf{z}_{t-1} = \frac{1}{\sqrt{1-\beta_t}}(\mathbf{z}_t - \frac{\beta_t}{\sqrt{1-\alpha_t^2}}\boldsymbol{\epsilon}_\theta(\mathbf{z}_t, t)) + \rho_t \boldsymbol{\epsilon}$
end for
$\mathbf{x}, \mathbf{h} \sim p_\xi(\mathbf{x}, \mathbf{h} | \mathbf{z}_{\mathrm{x},0}, \mathbf{z}_{\mathrm{h},0})$ {Decoding}
return $\mathbf{x}, \mathbf{h}$

- About the experiments, how to select an optimal value for T (timestep) to balance the model's performance and computational cost?

GEOLDM

- Why is Gaussian distribution chosen in the diffusion model, can it be effectively replaced by alternative distributions?

- How does latent modeling in the proposed method contribute to controllable molecule generation?