Contrastive Deep Fusion-based Diffusion Multi-view Drug Recommendation (Supplementary Material)

Anonymous Authors
Paper under double-blind review

Overview of the Supplementary Material

This document is part of the Supplementary Material (SM) for the following paper:

Contrastive Deep Fusion-based Diffusion Multi-view Drug Recommendation.

Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI-25), Montreal, Canada, 2025

The supplementary material provides implementation details for the representation diffusion in Section 3.3 Representation Diffusion Module.

Representation diffusion

We design forward and reverse processes for the diffusion of representations for diagnoses and procedures, separately. The forward process involves continuously adding incremental Gaussian noise to obtain complete Gaussian noise, while the reverse process employs neural networks for denoising to generate robust diagnosis and procedure representations. It is worth noting that we simplify the diffusion model as much as possible to emphasize the rationale and effectiveness of our proposed CD-MDRec in introducing the diffusion model into the field of drug recommendation.

Taking diagnoses as an example, we introduce the implementation process of representation diffusion.

① **Forward Process**: The process takes $E_d^{t,0}$ as the input representation of diffusion, and $E_d^{t,S}$ as the output representation of diffusion, where S denotes the total time steps. The transition process at step $s \in \{1, 2, \dots, S\}$ is parameterized as

$$q(E_d^{t,s}|E_d^{t,s-1}) = \mathcal{N}(E_d^{t,s}; \sqrt{1-\beta_s}E_d^{t,s-1}, \beta_s I) \qquad (1)$$

where $\beta_s = 1 - n_s/n_{s-1}$, $n_0 = 1$, $n_{1:S} = \{m_{max}, \dots, m_{min}\}$ forms a linearly decreasing set, with m_{max} and m_{min} being two hyperparameters set to default values of $1e-2$ and $1e-4$, respectively. Let $\overline{\alpha}_s = \prod_{s'=1}^{s}(1-\beta_{s'})$, then through iterative substitutions, the final $E_d^{t,s}$ can be simplified to

$$\begin{aligned} E_d^{t,s} &= \sqrt{1-\beta_s}E_d^{t,s-1} + \sqrt{\beta_s}\epsilon_1 \\ &= \sqrt{1-\beta_s}(\sqrt{1-\beta_{s-1}}E_d^{t,s-2} + \sqrt{\beta_{s-1}}\epsilon_2) + \sqrt{\beta_s}\epsilon_1 \\ &= \sqrt{1-\beta_s}\sqrt{1-\beta_{s-1}}E_d^{t,s-2} + \sqrt{1-(1-\beta_s)(1-\beta_{s-1})}\epsilon_3 \\ &= \sqrt{\overline{\alpha}_s}E_d^{t,0} + \sqrt{1-\overline{\alpha}_s}\epsilon \end{aligned} \qquad (2)$$

where ϵ, ϵ_1, ϵ_2 and ϵ_3 all follow Gaussian distributions with mean 0 and variance 1. Therefore, the diffusion model can directly obtain the diffusion results at step s through Eq. 2 without recursive calculations.

② **Reverse Process**: The process takes $E_d^{t,S}$ as the input representation, and $\widetilde{E}_d^{t,0}$ as the final output representation. The transition process at step $s \in \{1, 2, \dots, S\}$ is parameterized as

$$q(E_d^{t,s-1}|E_d^{t,s}) = \mathcal{N}(E_d^{t,s-1}; \mu_\theta(E_d^{t,s}, s), \sum_\theta(E_d^{t,s}, s)) \qquad (3)$$

where $\mu_\theta(\cdot)$ and $\sum_\theta(\cdot)$ are neural networks with learnable parameters θ. To simplify operations and enhance the model's generalization, we concatenate $E_d^{t,s}$ and h_s, followed by two layers of full connect neural network to obtain the final output representation $\widetilde{E}_d^{t,0}$, where h_s is the learnable mapping representation for step s.

③ **Diffusion Loss Function**: To enhance the generalization of the diffusion model, we employ a weighted Mean Square Error (MSE) loss as the loss function for the diffusion model of the t-th patient, as follows

$$L_{d,diff}^t = \frac{\overline{\alpha}_{s-1} - \overline{\alpha}_s}{2(1-\overline{\alpha}_{s-1})(1-\overline{\alpha}_s)}||E_d^{t,0} - \widetilde{E}_d^{t,0}||_2^2 \qquad (4)$$

where the time step s is obtained through random sampling. Therefore, the final diffusion loss is

$$L_{diff} = L_{diff}^1 + L_{diff}^2 + ... + L_{diff}^N \qquad (5)$$

where $L_{diff}^t = L_{d,diff}^t + L_{p,diff}^t$ of patient $t \in \{1, ..., N\}$.