04

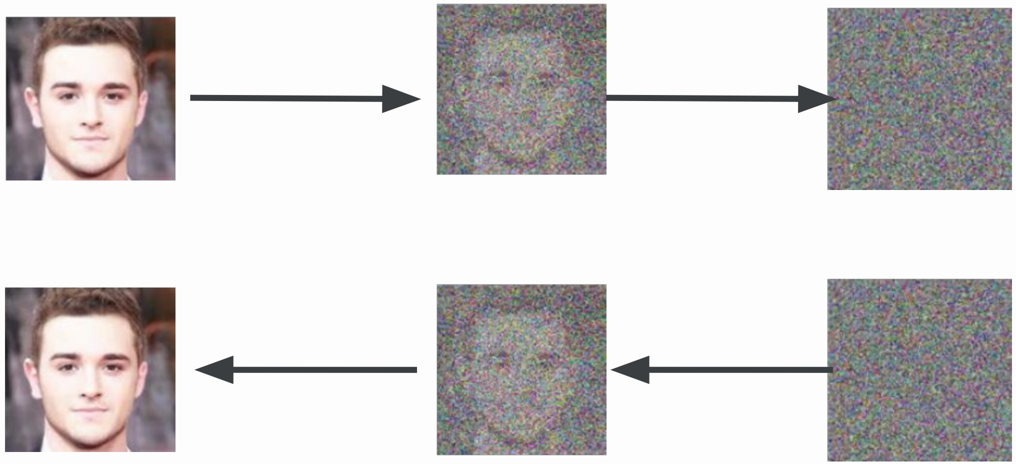
Diffusion Model

-based Generative Recommendation

What is Diffusion

Forward process

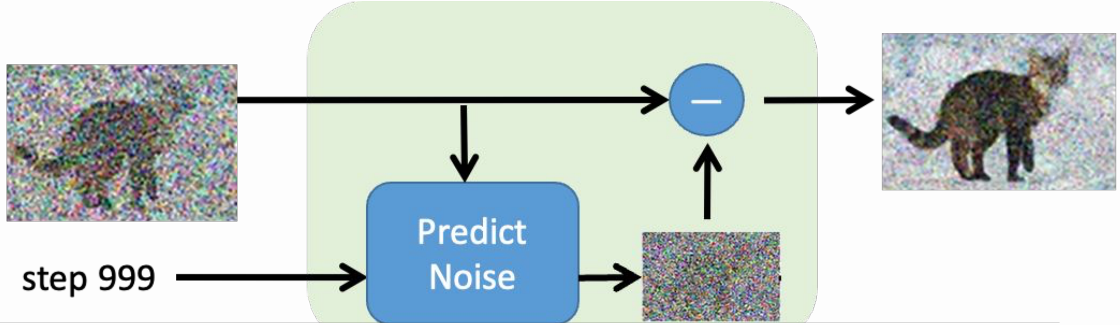
Reverse process



Build the mapping between data sample and Gaussian sample

Denoising Diffusion Probabilistic Models. NeurIPS 2020

What is Diffusion



Algorithm 1 Training

1: **repeat**
2: $\mathbf{x}_0 \sim q(\mathbf{x}_0)$
3: $t \sim \mathrm{Uniform}(\{1, \ldots, T\})$
4: $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$
5: Take gradient descent step on
$$\nabla_\theta \left\| \boldsymbol{\epsilon} - \boldsymbol{\epsilon}_\theta(\sqrt{\bar{\alpha}_t}\mathbf{x}_0 + \sqrt{1-\bar{\alpha}_t}\boldsymbol{\epsilon}, t) \right\|^2$$
6: **until** converged

Algorithm 2 Sampling

1: $\mathbf{x}_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$
2: **for** $t = T, \ldots, 1$ **do**
3: $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ if $t > 1$, else $\mathbf{z} = \mathbf{0}$
4: $\mathbf{x}_{t-1} = \frac{1}{\sqrt{\alpha_t}} \left(\mathbf{x}_t - \frac{1-\alpha_t}{\sqrt{1-\bar{\alpha}_t}} \boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t) \right) + \sigma_t \mathbf{z}$
5: **end for**
6: **return** $\mathbf{x}_0$

Remove the noise step by step from a Gaussian sample.

Denoising Diffusion Probabilistic Models. NeurIPS 2020

Diffusion in CV

Diffusion is at the core of visual content generation.

Image generation

Stable Diffusion, DALL-E...

Video generation

Sora, Hunyuan-Video, Keling...

Diffusion for recommendation

Use diffusion to enhance traditional recommender

- More robust representation
- Data augmentation

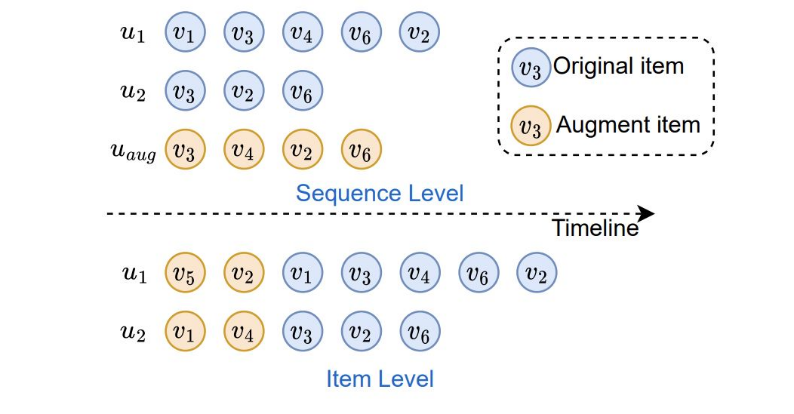
Diffusion as recommender

- Diffuse on the user interaction vector
- Diffuse on item representation
- Discrete diffusion

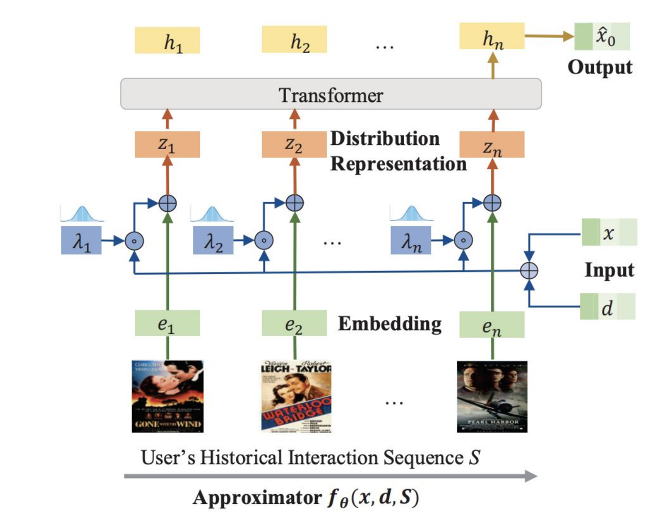
Diffusion for personalized content generation

- Personalized try-on, image,....

Diffusion as enhancer



Generate more interaction or sequences

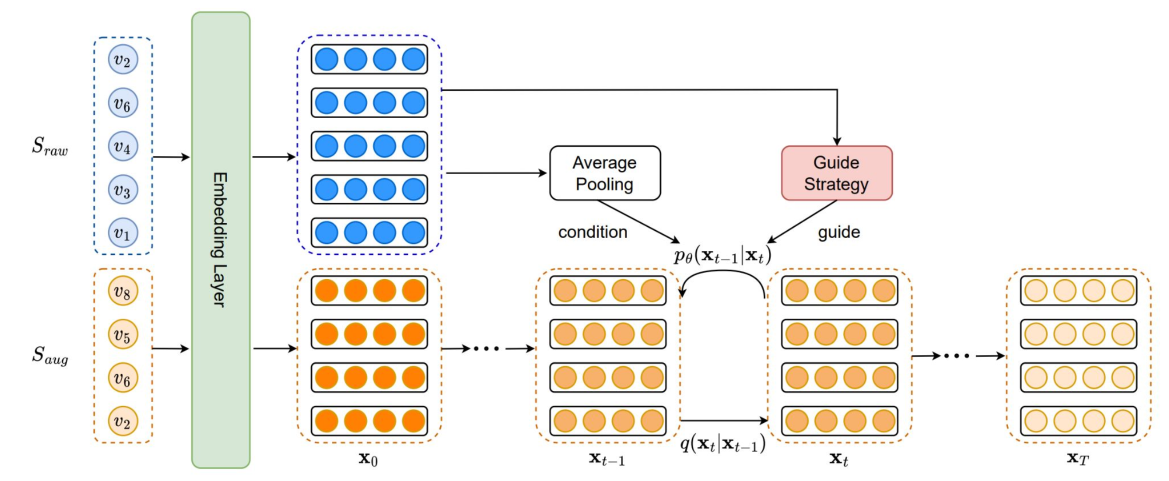


Enhance the robustness of embeddings

Diffusion Augmentation for Sequential Recommendation, in CIKM 2023.
DiffuRec: A Diffusion Model for Sequential Recommendation, in TOIS 2024

Pseudo sequence generation (I)

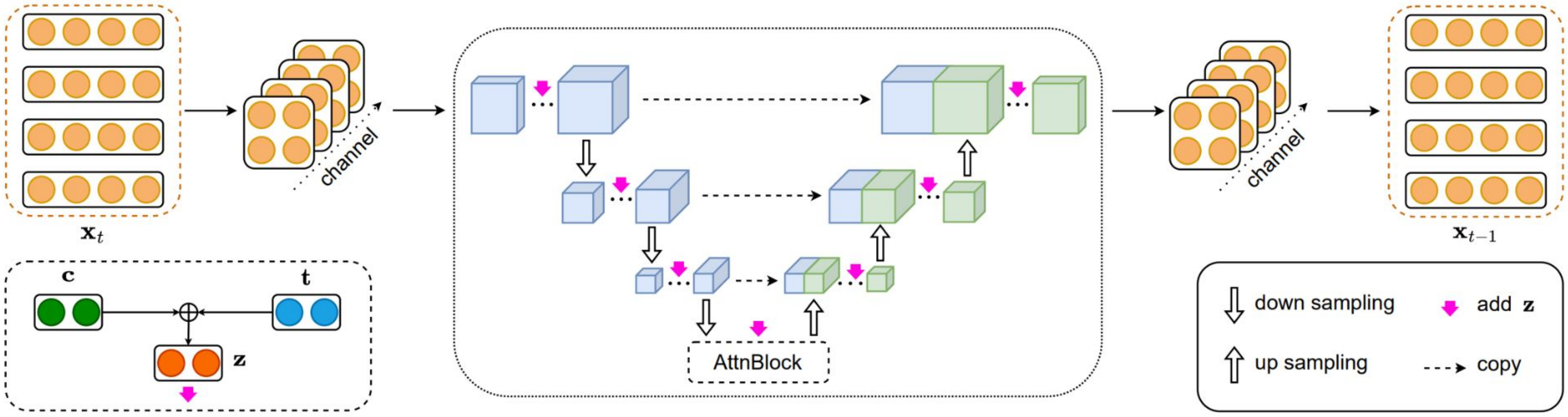
Generate pseudo sequence embeddings conditioned on historical interaction sequence



Diffusion Augmentation for Sequential Recommendation, in CIKM 2023.

Pseudo sequence generation (II)

The model architecture is adopted from U-Net



Diffusion Augmentation for Sequential Recommendation, in CIKM 2023.

Diffusion for recommendation

Use diffusion to enhance traditional recommender

- More robust representation
- Data augmentation

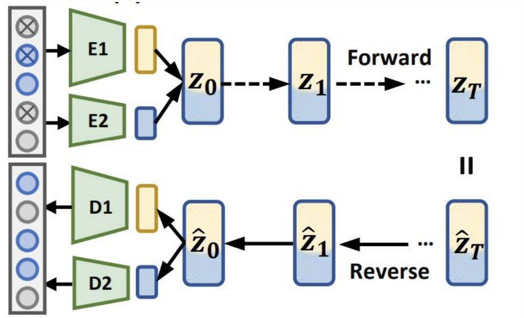
Diffusion as recommender

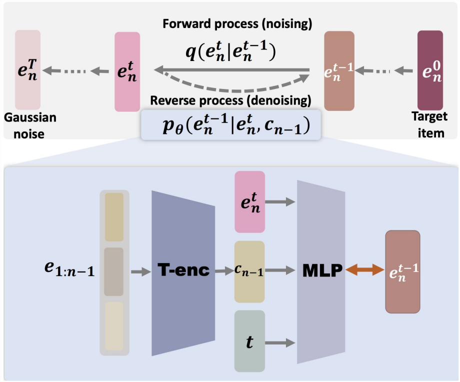
- Diffuse on the user interaction vector
- Diffuse on item representation
- Discrete diffusion

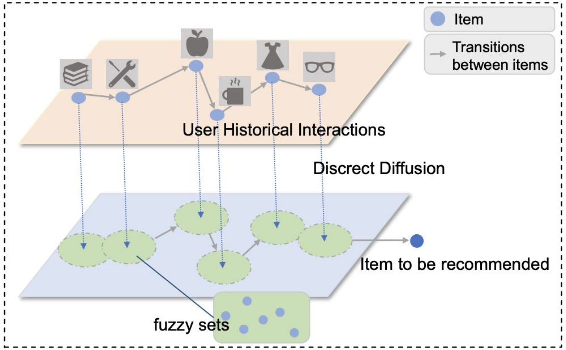
Diffusion for personalized content generation

- Personalized try-on, image,....

Diffusion as recommender







Diffuse on the user interaction vector

Diffuse on item representation

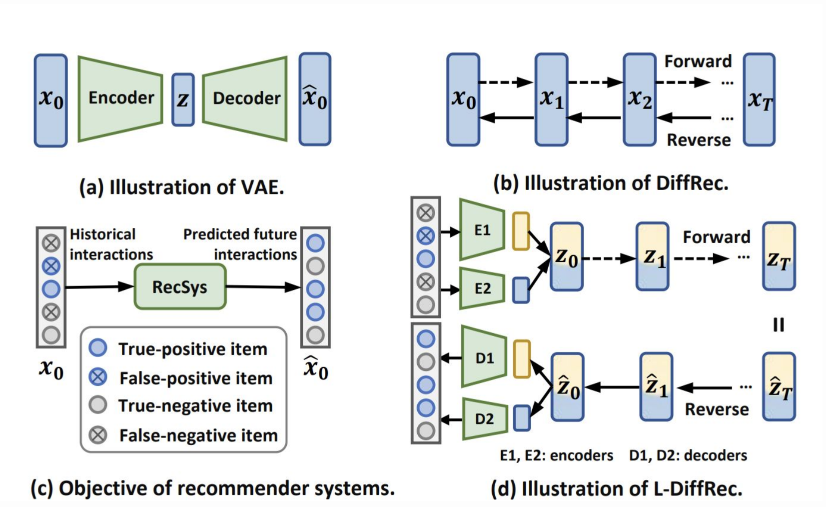
Discrete diffusion

Diffusion Recommender Model, in SIGIR 2023.
Generate What You Prefer: Reshaping Sequential Recommendation via Guided Diffusion, in NeurIPS 2023
Breaking Determinism: Fuzzy Modeling of Sequential Recommendation Using Discrete State Space Diffusion Model, in NeurIPS 2024

Interaction vector completion (I)

Motivation – limitation of GANs and VAEs:

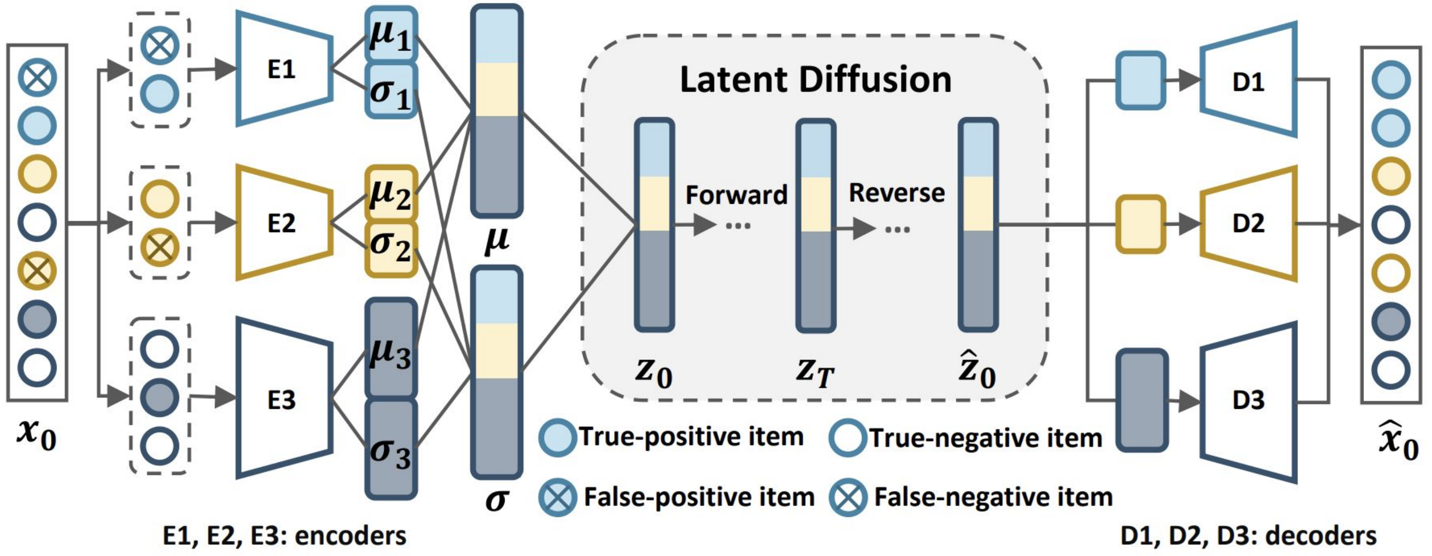
GAN- and VAE-based recommenders suffers from issues like **instability and representation collapse**.



Diffusion Recommender Model, in SIGIR 2023.

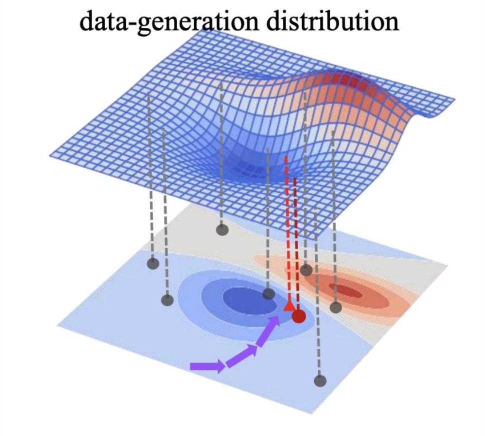
Interaction vector completion (II)

Forward: **corrupt the interaction vector** into gaussian noise
Reverse: **recover the interaction vector** from the gaussian



Generate item embedding

There exists an implicit distribution, from which target item embedding can be generated.



Learning-to-generate Paradigm

Challenge:

- The data-generation distribution is complicated and unknown.

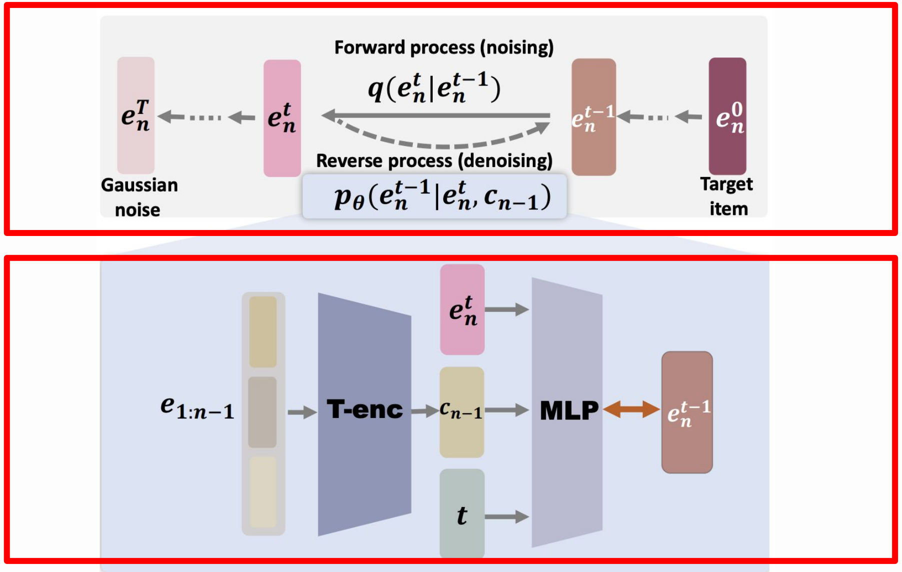
Solution:

- Capture the data-generation distribution by connecting it with Gaussian distribution.
- This can be achieved by diffusion.

Generate What You Prefer: Reshaping Sequential Recommendation via Guided Diffusion, in NeurIPS 2023

Generate item embedding

- Diffusion on target item embeddings.
- Guided by user interaction sequence for personalization.



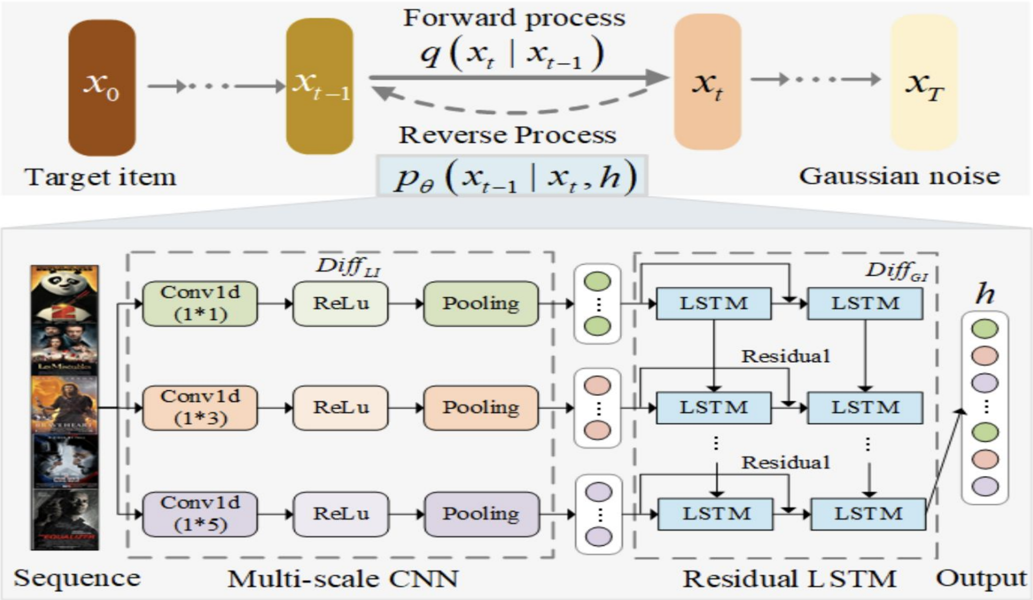
Diffusion process

Guidance

Generate What You Prefer: Reshaping Sequential Recommendation via Guided Diffusion, in NeurIPS 2023

Generate item embedding

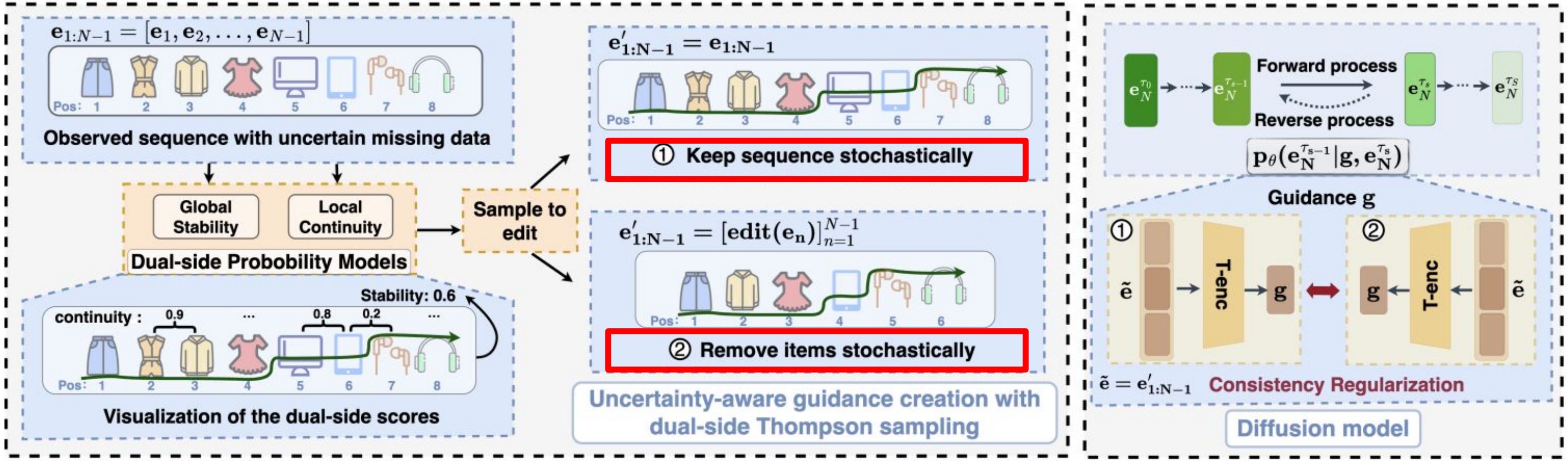
- Different sequence encoder



Diffusion Recommendation with Implicit Sequence Influence, in WWW 2024

Generate item embedding

- Uncertainty-aware guidance



Uncertainty-aware Guided Diffusion for Missing Data in Sequential Recommendation, in SIGIR 2025

Generate item embedding

- Incorporate preference optimization

$$\mathcal{L}_{\text{Simple}} = \mathbb{E}_{(\mathbf{e}_0^+, \mathbf{c}, t)} \left[\left\| \mathcal{F}_\theta(\mathbf{e}_t^+, t, \mathcal{M}(\mathbf{c})) - \mathbf{e}_0^+ \right\|_2^2 \right] ,$$

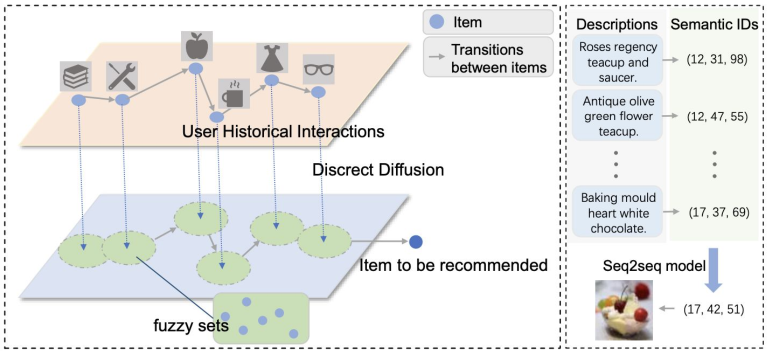
$$\mathcal{L}_{\text{BPR-Diff-C}} = -\log \sigma(-|\mathcal{H}| \cdot [S(\hat{\mathbf{e}}_0^+, \mathbf{e}_0^+) - S(\mathcal{F}_\theta(\bar{\mathbf{e}}_t^-, t, \mathcal{M}(\mathbf{c})), \bar{\mathbf{e}}_0^-)]) .$$

$$\mathcal{L}_{\text{PerferDiff}} = \underbrace{\lambda \mathcal{L}_{\text{Simple}}}_{\text{Learning Generation}} + \underbrace{(1-\lambda)\mathcal{L}_{\text{BPR-Diff-C}}}_{\text{Learning Preference}} .$$

Preference Diffusion for Recommendation, in ICLR 2025

Discrete diffusion

State transitions occur under discrete conditions for the entire interaction sequence.



- Represent interaction sequence as one-hot vector through semantic ID.
- Conduct discrete diffusion on interaction sequence.

Breaking Determinism: Fuzzy Modeling of Sequential Recommendation Using Discrete State Space Diffusion Model, in NeurIPS 2024.

Discrete diffusion

Semantic IDs

Algorithm 1 Training of DDSR.

Input: historical interaction sequence $v_{1:n-1} = c_{1:n-1;1:m}$; target item $v_n = c_{n;1:m}$; transition matrix $\boldsymbol{Q}_t$; Approximator $f_\theta(\cdot)$.

Output: well-trained Approximator $f_\theta(\cdot)$.

While not converged do:

Sample Diffusion Time: $t \sim [0, 1, \ldots, T]$;

Calculate t-step transition probability: $\overline{\boldsymbol{Q}_t} = \boldsymbol{Q}_1\boldsymbol{Q}_2 \cdots \boldsymbol{Q}_t$;

Convert $c_{n;1:m}$ to one-hot encoding $\boldsymbol{x}^0_{n;1:m}$;

Obtain the discrete state $x^t_{n;1:m}$ after t steps by Equation 2, thereby obtaining the 'fuzzy set' $c^t_{1:n-1;1:m}$;

Modeling $c_{2;n;1:m}$ based on 'fuzzy sets' through Equation 5;

Take gradient descent step on $\nabla L_{CE}\left(\hat{c}_{2:n;1:m}, c_{2:n;1:m}\right)$.

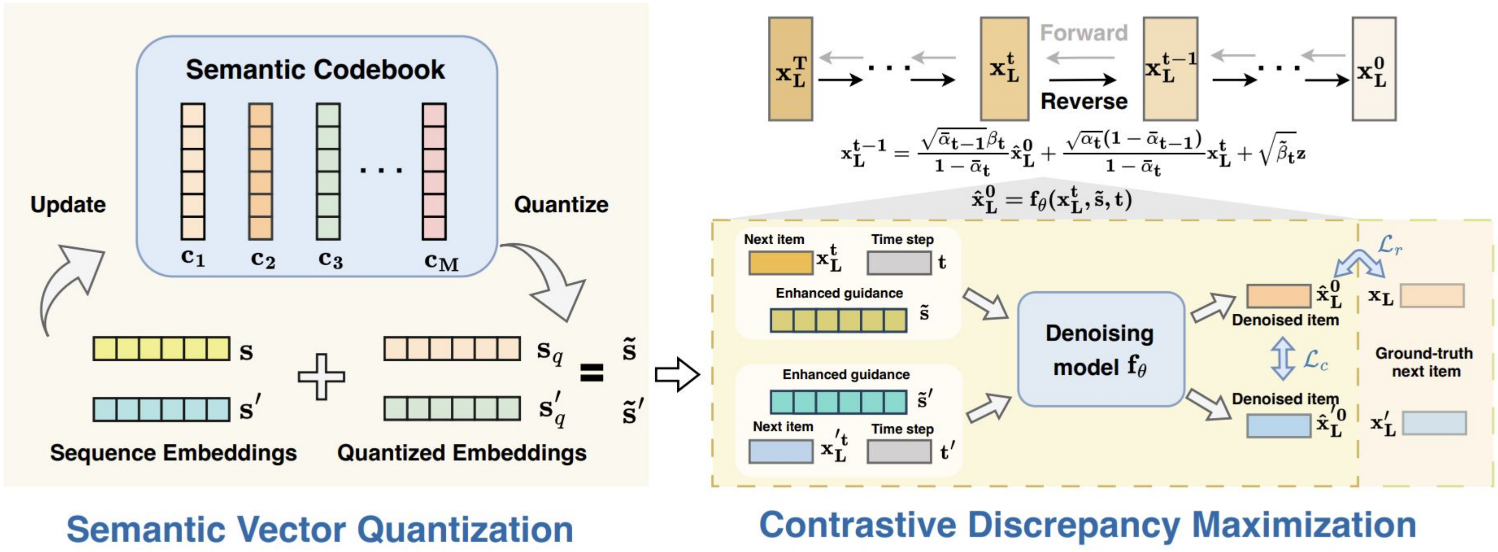
Forward process

$$[\boldsymbol{Q}_t]_{ij} = \begin{cases} (1-\beta_t)/(|\mathcal{V}|-1) & \text{if } i \neq j \\ \beta_t & \text{if } i = j \end{cases}.$$

$$\hat{c}_{2:n;1:m} = f_\theta(c^t_{1:n-1;1:m}, t).$$

Discrete diffusion

- Quantization embedding with continuous diffusion.



Distinguished Quantized Guidance for Diffusion-based Sequence Recommendation, in WWW 2025..

Diffusion for recommendation

Use diffusion to enhance traditional recommender

- More robust representation
- Data augmentation

Diffusion as recommender

- Diffuse on the user interaction vector
- Diffuse on item representation
- Discrete diffusion

Diffusion for personalized content generation

- Personalized try-on, image,....

Personalized content generation

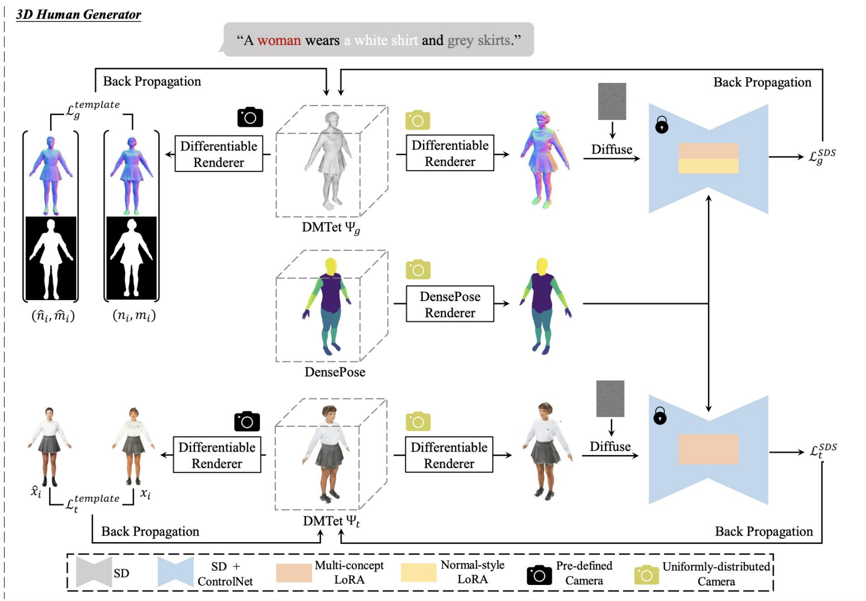
Personalized try-on

Personalized image

OOTDiffusion: Outfitting Fusion based Latent Diffusion for Controllable Virtual Try-on, in arXiv
InstantBooth: Personalized Text-to-Image Generation without Test-Time Finetuning. In CVPR 2024.

Personalized Try-on

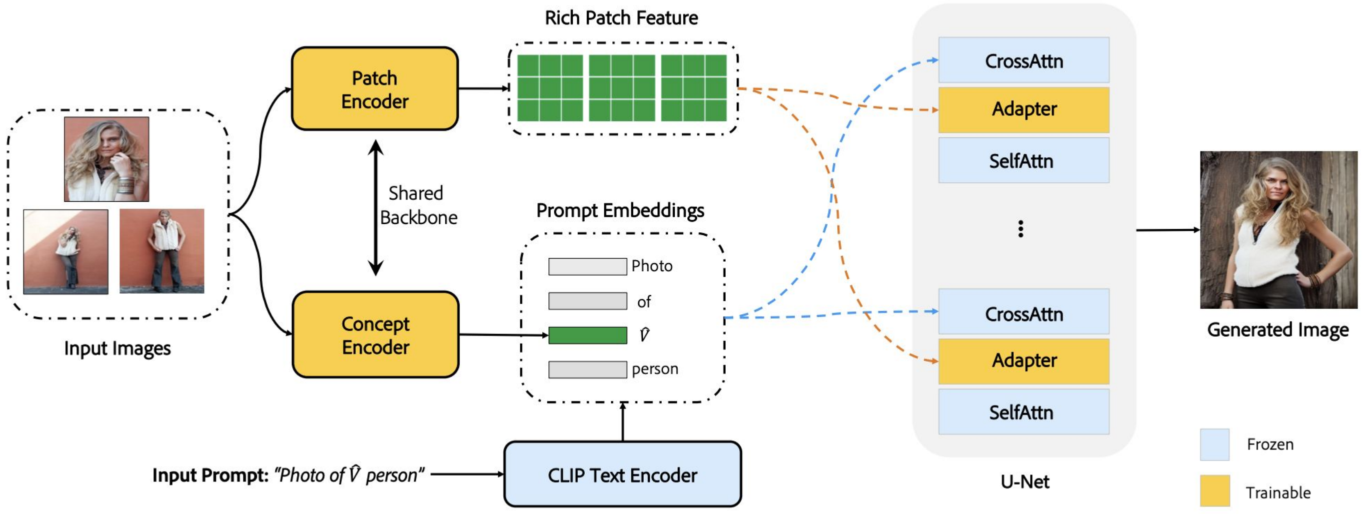
Generate realistic 3D try-on given person images, clothes images, and a text prompt.



DreamVTON: Customizing 3D Virtual Try-on with Personalized Diffusion Models. MM 2024

Personalized Image

Generate personalized image given person images and the desired concept.



InstantBooth: Personalized Text-to-Image Generation without Test-Time Finetuning. CVPR 2024