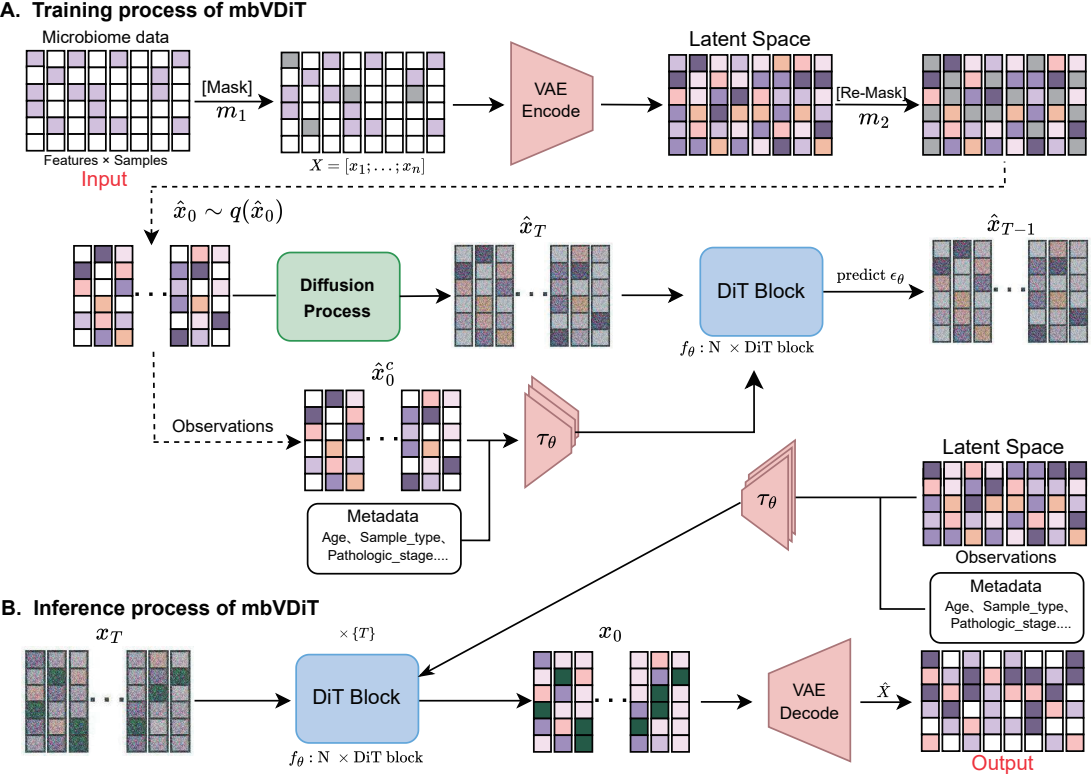
A. Training process of mbVDiT
Microbiome data
Features × Samples
Input
[Mask]
m_1
$X = [x_1; \ldots; x_n]$
VAE Encode
Latent Space
[Re-Mask]
m_2
$\hat{x}_0 \sim q(\hat{x}_0)$
Diffusion Process
$\hat{x}_T$
DiT Block
f_θ : N × DiT block
predict ϵ_θ
$\hat{x}_{T-1}$
Observations
$\hat{x}_0^c$
τ_θ
Metadata
Age、Sample_type、Pathologic_stage....
Latent Space
Observations
Metadata
Age、Sample_type、Pathologic_stage....
B. Inference process of mbVDiT
x_T
× {T}
DiT Block
f_θ : N × DiT block
x_0
VAE Decode
$\hat{X}$
Output