Diffusion Models (Forward Process)

Dr. Alireza Aghamohammadi

Encoder

- ❖ Let x be an image from the training set. Blend it with Gaussian noise independently for each pixel to get a noisy image.
- ❖ The forward diffusion process transforms the data example x through a series of intermediate variables $z_1, z_2, \ldots, z_T$ with the same size as x as follows:

$$
\begin{aligned}
z_1 &= \sqrt{1-\beta_1} \cdot x + \sqrt{\beta_1} \cdot \epsilon_1 \\
z_t &= \sqrt{1-\beta_t} \cdot z_{t-1} + \sqrt{\beta_t} \cdot \epsilon_t \quad \forall t \in \{2, \ldots, T\}
\end{aligned}
$$

  Here, $\epsilon_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$ represents Gaussian noise, and the probability distribution of z_t depends only on z_{t-1}.

- ❖ The hyperparameters $\beta_t \in [0, 1]$, known as the noise schedule, control the rate of noise blending.
- ❖ This ensures that the mean of z_t is closer to zero than the mean of z_{t-1}, and the variance of z_t is closer to the identity matrix than the variance of z_{t-1}.
- ❖ The transformations can be written as:

$$
\begin{aligned}
q(z_1 \mid x) &= \mathcal{N}(z_1; \sqrt{1-\beta_1}x, \beta_1 \boldsymbol{I}) \\
q(z_t \mid z_{t-1}) &= \mathcal{N}(z_t; \sqrt{1-\beta_t}z_{t-1}, \beta_t \boldsymbol{I})
\end{aligned}
$$