

Практика

Диффузионные модели

Туем Гислен

11-04-2024

Содержание

1	Введение	5
2	Теоретические основы	6
2.1	Определение	6
3	Пример генерации в PyTorch	8
4	Заключение	9

Список иллюстраций

Список таблиц

1 Введение

****Актуальность темы**:**

Диффузионные модели революционизировали генеративный ИИ, превзойдя GAN в качестве синт. Наряду с GAN, модели диффузии (diffusion models) или диффузионные модели являются одними из самых популярных.

Многие из основных идей, лежащих в основе диффузионных моделей, имеют сходство с более ранними моделями.

Основная идея диффузионных моделей основана на наблюдении, что шум трудно преобразовать в зашумленную версию x_0 , пропуская данные через N шагов стохастического кодировщика $q(x_N|x_0)$. После достаточного количества шагов мы (приблизительно) получаем стандартное многомерное или какое-либо другое удобное известное распределение. Затем мы изучаем обратный процесс $p(x_0|x_N)$ шагов декодера $p(x_{t-1}|x_t)$, пока мы не сгенерируем первоначальные данные x_0 .

Цель работы:

Систематизировать математические основы, архитектурные решения и прикладные аспекты диффузионных моделей.

2 Теоретические основы

2.1 Определение

Диффузионные модели – это генеративные модели, обучающиеся через последовательное зашумление

$$q(x_t|x_{t-1}) = \mathcal{N}(x_t; \sqrt{1-\beta_t}x_{t-1}, \beta_t\mathbf{I})$$

Связь с другими методами

Метод	Отличия
VAE	Используют энкодер-декодер
GAN	Опираются на состязательность
Score-Based	Общие теоретические корни

Принцип работы

Прямой процесс

1. Итеративное добавление гауссова шума
2. Марковская цепь с параметрами β_t

Обратный процесс

$$p_{\theta}(x_{t-1}|x_t) = \mathcal{N}(\mu_{\theta}(x_t, t), \Sigma_{\theta}(x_t, t))$$

Применение

3 Пример генерации в PyTorch

```
from diffusers import StableDiffusionPipeline pipe = StableDiffusionPipeline.from_pretrained("stablediffusion-2") image = pipe("кошка в шляпе").images[0] “
```


4 Заключение

Диффузионные модели демонстрируют: 1. Более стабильное обучение vs GAN
2. Высокое качество генерации 3. Широкие перспективы в мультимодальных приложениях