

1- VAEs:

سوالات تشریحی:

الف) تفاوت اصلی بین یک اتوانکودر استاندارد و یک وریشنال اتوانکودر چیست؟ چرا VAE از توزیع احتمال در فضای نهان¹ استفاده می‌کند؟

ب) اجزای مختلف تابع خطا در VAE را بیان کنید و توضیح دهید که چرا استفاده از آن‌ها ضروری است؟
ج) چرا VAE به جای نگاشت داده‌ها به نقاط ثابت در فضای نهان، آن‌ها را به توزیع‌هایی مانند گاوسی نگاشت می‌کند؟

پیاده‌سازی:

یک VAE ساده برای بازسازی مجموعه داده MNIST پیاده‌سازی کنید. بخش‌های انکودر و دیکودر را طراحی کرده و عناصر موجود در تابع خطا را مشخص کنید. اندازه بردار نهان را ۲ در نظر بگیرید. در انتها کیفیت تصاویر بازسازی شده را با معیارهای مناسب بسنجید.

سپس پس از آموزش کامل مدل، فضای نهان را برای نمونه‌های تصادفی از مجموعه داده رسم کنید. اگر اندازه بردار نهان ۲ باشد، نقاط در فضای دوبعدی نمایش داده شوند.

امتیازی ابعاد فضای نهان را به ۴ و ۱۶ تغییر بدهید و کیفیت تصاویر بازسازی شده در حالت‌هایی با اندازه بردار نهان ۲، ۴ و ۱۶ را با هم مقایسه کنید. برای ترسیم فضای نهان برای بردارهای نهان با اندازه ۴ و ۱۶ از T-SNE استفاده کنید تا ابعاد را به ۲ کاهش دهید تا قابل ترسیم باشد.

2- GANs:

سوالات تشریحی:

الف) معماری یک GAN را توضیح دهید. نقش هر یک از اجزای Generator و Discriminator چیست؟
ب) تابع زیان GAN چگونه تعریف می‌شود؟ چرا این تابع به صورت رقابتی طراحی شده است؟
ج) برخی از پیشرفت‌های اخیر در معماری‌های GANها (DCGAN، StyleGAN، CycleGAN) را توضیح دهید. چه تفاوت‌هایی بین این مدل‌ها و GAN استاندارد وجود دارد؟
د) شباهت‌ها و تفاوت‌های GAN و VAE را توضیح دهید. در چه شرایطی یکی از این مدل‌ها نسبت به دیگری ترجیح داده می‌شود؟

سوال پیاده‌سازی:

سوال مربوط به رفع نویز در تمرین قبلی را این بار با معماری GAN پاسخ دهید.

3- Diffusion Models:

سوالات تشریحی:

الف) ایده اصلی این معماری چیست و چگونه عمل می‌کند؟
ب) فرآیند انتشار (Forward Process) در مدل‌های دیفیوژن چگونه عمل می‌کند؟ چرا نویز گاوسی به داده‌ها اضافه می‌شود؟ توسط چه تابعی و چگونه نویز به تصاویر اضافه می‌شود؟
پ) فرآیند معکوس (Reverse Process) در مدل‌های دیفیوژن چگونه داده‌های نویزی را به داده‌های اصلی بازسازی می‌کند؟

¹ Latent Space

ت) تابع خطا در مدل‌های دیفیوژن چگونه تعریف می‌شود؟ چرا یادگیری پارامترهای فرآیند معکوس مهم است؟
ث) چه معماری‌هایی معمولاً برای مدل‌های دیفیوژن استفاده می‌شوند؟ چگونه شبکه‌های عصبی برای پیش‌بینی نویز یا توزیع داده‌ها آموزش داده می‌شوند؟
ج) چگونه تعداد مراحل دیفیوژن بر کیفیت بازسازی داده‌ها و سرعت فرآیند تأثیر می‌گذارد؟
چ) برخی از مزایا و معایب مدل‌های دیفیوژن نسبت به GAN ها و VAE ها را بیان کنید.

سوال پیاده‌سازی:

یک مدل دیفیوژن شرطی (Conditional Diffusion Model) برای مجموعه داده CIFAR-10 پیاده‌سازی کنید. این مدل باید بتواند فرآیند انتشار را با افزودن نویز گوسی به تصاویر اصلی انجام دهد و در فرآیند معکوس، از تصاویر نویزی شروع کرده و داده‌های اصلی را بازسازی کند. هدف این است که با شرطی‌سازی مدل روی کلاس‌های CIFAR-10 (گربه، هواپیما، یا ...)، از نویز کاملاً تصادفی به یک نمونه واضح و باکیفیت از کلاس اعلام‌شده برسیم. در ابتدا فرآیند انتشار را پیاده‌سازی کرده و تصاویر نویزی در مراحل مختلف را تجسم کنید. سپس یک شبکه عصبی برای فرآیند معکوس طراحی کنید که بتواند نویز را حذف کند و تصاویر را بازسازی نماید. مدل خود را برای تولید تصاویر جدید از کلاس‌های خاص آزمایش کنید و کیفیت بازسازی را برای کلاس‌های مختلف تحلیل کنید. در نهایت چند تصویر از کلاس‌های اعلام شده از مدل خروجی بگیرید و آن‌ها را نمایش دهید.

امتیازی) تأثیر تعداد مراحل و پارامترهای نویز را بر کیفیت بازسازی بررسی کرده و نتایج را به صورت نمودار و جدول گزارش دهید.

نکات :

فرمت فایل pdf به این صورت باشد :

HW3-Full Name-Student ID

- تصاویر دیتاست در فایل **زیپ** موجود است.
- لینک کولب در گزارش آورده شود و تمام سلول‌ها **اجرا شده** و **قابل دسترسی** باشند.