



خلاصه دوره‌های هوش مصنوعی گوگل – یادداشت مطالعه به فارسی

آشنایی با تولید تصویر

Introduction to Image Generation

ارائه‌دهنده: Google Cloud / Google Skills

دوره ریزآموزشی – ۳۰ دقیقه

سطح: میانی – رایگان با نشان

فایل ۱۶ از ۲۲ – تهیه شده برای تدریس

خلاصه کلی دوره

کوتاه اما مفهومی. توضیح می‌دهد مدل‌های انتشار – خانواده مدلی که پشت اکثر تولیدکننده‌های تصویر امروزی است – چطور کار می‌کنند، و ایده اصلی‌شان از کجا آمده: فیزیک و ترمودینامیک.

دوره از مفهوم به عمل می‌رود: بعد از توضیح نظری، آموزش و استقرار مدل روی Vertex AI را نشان می‌دهد.

نکات مهم هر بخش

۱. جایگاه مدل‌های انتشار

- خانواده‌های مدل‌های مولد تصویر: مدل‌های مبتنی بر GAN، VAE، مدل‌های اتورگرسیو، و مدل‌های انتشار.
- مدل‌های انتشار در سال‌های اخیر به دلیل کیفیت و پایداری آموزش، غالب شده‌اند.

۲. ایده اصلی: الهام از ترمودینامیک

- تشبیه به پخش شدن قطره جوهر در آب: فرایندی که به‌طور طبیعی به بی‌نظمی می‌رود.
- سؤال کلیدی: اگر بشود این فرایند را معکوس کرد، می‌توان از بی‌نظمی به نظم رسید.

۳. فرایند رو به جلو (Forward diffusion)

- به تصویر واقعی، گام‌به‌گام نویز اضافه می‌شود تا در نهایت به نویز خالص تبدیل شود.
- این مرحله محاسبات ساده‌ای دارد و برای ساختن داده آموزشی استفاده می‌شود.

۴. فرایند معکوس (Reverse denoising)

- مدل یاد می‌گیرد در هر گام نویز اضافه‌شده را پیش‌بینی و حذف کند.
- در زمان تولید، از نویز تصادفی شروع می‌کنیم و مدل گام‌به‌گام آن را به تصویر تبدیل می‌کند.
- افزودن شرط متنی (prompt) به این فرایند، همان چیزی است که «تولید تصویر از متن» را ممکن می‌کند.

۵. کاربردها و ابزارها

- ویرایش تصویر، پر کردن نواحی حذف شده (inpainting)، گسترش تصویر (outpainting)، افزایش وضوح.
- خانواده Imagen گوگل و در دسترس بودن آن روی Vertex AI.
- آموزش و استقرار مدل روی Vertex AI و ملاحظات هزینه و مقیاس.

۶. ملاحظات مسئولانه

- خطر تولید محتوای جعلی و دیپ فیک.
- سوگیری در بازنمایی افراد و فرهنگ‌ها در تصاویر تولیدشده.
- مالکیت فکری و منشأ داده آموزشی.
- ابزارهای نشان‌گذاری (watermarking) برای تشخیص محتوای تولیدشده توسط AI.

جمع‌بندی سریع برای مرور

جمله‌ای که همه چیز را خلاصه می‌کند

مدل انتشار یاد می‌گیرد نویز را حذف کند؛ تولید تصویر یعنی همین حذف نویز را از یک نویز خالص شروع کنیم و با متن هدایتش کنیم.

- ایده مرکزی: افزودن نویز آسان است، حذف آن هنر است – و مدل دقیقاً همان را یاد می‌گیرد.
- کیفیت خروجی به سه چیز بستگی دارد: مدل، پرامپت، و تعداد گام‌های حذف نویز.
- این دوره بخشی از مسیر پیشرفته Generative AI for Developers است.

منابع

- https://www.skills.google/paths/183/course_templates/541
- <https://cloud.google.com/blog/topics/training-certifications/new-generative-ai-trainings-from-google-cloud>