

فهرست مطالب

پیشگفتار.....	۹
درباره این کتاب.....	۱۱
چه کسانی باید این کتاب را بخوانند.....	۱۳
نحوه سازماندهی این کتاب: یک نقشه راه.....	۱۴
درباره کد.....	۱۵
سایر منابع آنلاین.....	۱۶
در مورد نویسنده.....	۱۷
در مورد تصویر روی جلد.....	۱۹
فصل اول: درک مدل‌های زبانی بزرگ.....	۲۱
۱-۱ مدل زبانی بزرگ چیست؟.....	۲۳
۱-۲ کاربردهای مدل‌های زبانی بزرگ.....	۲۵
۱-۳ مراحل ساخت و استفاده از مدل زبانی بزرگ.....	۲۷
۱-۴ معرفی معماری ترنسفورمر.....	۲۹
۱-۵ استفاده از مجموعه داده‌های بزرگ.....	۳۴
۱-۶ نگاهی دقیق‌تر به معماری GPT.....	۳۶
۱-۷ ساخت یک مدل زبانی بزرگ.....	۳۸
خلاصه.....	۳۹
فصل دوم: کار با داده‌های متنی.....	۴۱
۲-۱ درک تعبیه‌های کلمه.....	۴۲
۲-۲ توکن‌سازی متن.....	۴۵

۵۰	۲-۳ تبدیل توکن‌ها به شناسه‌های توکن
۵۵	۲-۴ افزودن توکن‌های زمینه خاص
۵۹	۲-۵ رمزگذاری زوج بایت
۶۲	۲-۶ نمونه برداری داده با یک پنجره کشویی
۷۰	۲-۷ ایجاد تعبیه‌های توکن
۷۳	۲-۸ رمزگذاری موقعیت کلمات
۷۸	خلاصه

فصل سوم: کدنویسی سازوکارهای توجه

۸۱	۳-۱ مشکل مدل سازی دنباله‌های طولانی
۸۵	۳-۲ ثبت وابستگی‌های داده با سازوکارهای توجه
۸۶	۳-۳ توجه به بخش‌های مختلف ورودی با خودتوجه
۹۷	۳-۴ پیاده‌سازی خودتوجهی با وزن‌های آموزش پذیر
۱۰۹	۳-۵ پنهان کردن کلمات آینده با توجه علی
۱۱۹	۳-۶ گسترش توجه تک‌سر به توجه چندسر
۱۳۰	خلاصه

فصل چهارم: پیاده‌سازی یک مدل GPT

۱۳۱	برای تولید متن (از پایه)
۱۳۲	۴-۱ کدنویسی معماری مدل زبانی بزرگ
۱۴۰	۴-۲ نرمال‌سازی فعال‌سازها با نرمال‌سازی لایه
۱۴۷	۴-۳ پیاده‌سازی یک شبکه پیشخور با فعال‌سازی‌های GELU
۱۵۲	۴-۴ افزودن اتصالات میانبر
۱۵۶	۴-۵ اتصال لایه‌های توجه و لایه خطی در یک بلوک ترنسفورمر
۱۶۰	۴-۶ کدنویسی مدل GPT
۱۶۷	۴-۷ تولید متن
۱۷۲	خلاصه

فصل پنجم: پیش‌آموزش روی داده‌های بدون برچسب

۱۷۶	۵-۱ ارزیابی مدل‌های متن مولد
-----	------------------------------

۱۹۵	۵-۲ آموزش یک مدل زبانی بزرگ
۲۰۲	۵-۳ استراتژی‌های رمزگشایی برای کنترل تصادفی بودن
۲۱۱	۵-۴ بارگذاری و ذخیره‌سازی وزن‌های مدل در PyTorch
۲۱۳	۵-۵ بارگذاری وزن‌های از پیش آموزش داده‌شده از OpenAI
۲۲۲	خلاصه

فصل ششم: تنظیم دقیق برای طبقه‌بندی ۲۲۳

۲۲۴	۶-۱ روش‌های مختلف تنظیم دقیق
۲۲۶	۶-۲ آماده‌سازی مجموعه داده
۲۳۱	۶-۳ ایجاد بارگذارهای داده
۲۳۷	۶-۴ مقداردهی اولیه مدل با وزن‌های از پیش آموزش داده شده
۲۴۰	۶-۵ اضافه کردن یک سر طبقه‌بندی
۲۴۸	۶-۶ محاسبه خطا و درستی طبقه‌بندی
۲۵۳	۶-۷ تنظیم دقیق مدل روی داده‌های نظارت‌شده
۲۶۰	۶-۸ استفاده از مدل زبانی بزرگ به‌عنوان یک طبقه‌بند هرزنامه
۲۶۲	خلاصه

فصل هفتم: تنظیم دقیق برای پیروی کردن از دستورعمل‌ها ۲۶۵

۲۶۶	۷-۱ مقدمه‌ای بر تنظیم دقیق دستورعمل
۲۶۷	۷-۲ آماده‌سازی مجموعه داده برای تنظیم دقیق دستورعمل نظارت‌شده
۲۷۳	۷-۳ سازماندهی داده‌ها در دسته‌های آموزشی
۲۸۷	۷-۴ ایجاد بارگذارهای داده برای یک مجموعه داده دستورعمل
۲۹۰	۷-۵ بارگذاری یک مدل زبانی بزرگ از پیش آموزش دیده
۲۹۴	۷-۶ بهینه‌سازی مدل زبانی بزرگ روی داده‌های دستورعمل
۲۹۹	۷-۷ استخراج و ذخیره پاسخ‌ها
۳۰۵	۷-۸ ارزیابی مدل زبانی بزرگ به‌دقت تنظیم‌شده
۳۱۶	۷-۹ نتیجه‌گیری
۳۱۸	خلاصه

پیوست الف: مقدمه‌ای بر PyTorch ۳۱۹

الف-۱ PyTorch چیست؟ ۳۱۹

الف-۲ درک تانسورها ۳۲۷

الف-۳ مدل‌ها به عنوان نمودارهای محاسباتی ۳۳۲

الف-۴ مشتق‌گیری خودکار آسان می‌شود ۳۳۳

الف-۵ پیاده‌سازی شبکه‌های عصبی چندلایه ۳۳۶

الف-۶ راه‌اندازی بارگذارهای داده کارآمد ۳۴۲

الف-۷ یک حلقه آموزشی معمولی ۳۴۸

الف-۸ ذخیره و بارگذاری مدل‌ها ۳۵۳

الف-۹ بهینه‌سازی عملکرد آموزش با GPU ها ۳۵۴

خلاصه ۳۶۵

پیوست ب: منابع و مطالعه بیشتر ۳۶۷

پیوست د: افزودن ویژگی‌های تکمیلی به حلقه آموزش ۳۹۷

د-۱ نرخ یادگیری ۳۹۹

د-۲ کاهش کسینوسی ۴۰۱

د-۳ برش گرادیان ۴۰۲

د-۴ تابع آموزشی اصلاح‌شده ۴۰۴

پیوست ه: تنظیم دقیق با بهره‌وری پارامتر با استفاده از LoRA ۴۰۹

ه-۱ مقدمه‌ای بر LoRA ۴۰۹

ه-۲ آماده‌سازی مجموعه داده ۴۱۱

ه-۳ مقداردهی اولیه مدل ۴۱۴

ه-۴ تنظیم دقیق با کارایی پارامتر بالا با استفاده از LoRA ۴۱۶