

پردازش زبان‌های طبیعی – تمرین چهارم

گزارش کار – مهرشاد فلاح اسطلخ‌زیر

401521462

منطق کد:

ابتدا کتابخانه‌هایی که در colab موجود نیست را نصب می‌کنم. در مرحله بعد توابع و کتابخانه‌های مورد نیاز را import می‌کنم. با استفاده از huggingface_hub و تابع login و access token که در پروفایل شخصی خودم گرفته‌ام به huggingface ورود می‌کنم. در مرحله بعدی و با استفاده از تابع جمی‌گذارم و گزینه درست را هم همراه آن به تمام گزینه‌ها اضافه می‌کنم و سوال را هم به این صورت می‌گذارم که کدام بیماری تشخیص درست برای علائم گفته شده در متن است. بعد از آن دیتای اصلی را که شبیه یک json است وارد می‌کنم (تمام این دیتا رندوم است چرا که ChatGPT قادر به خواندن PDF ارسال شده نبود و مجبور شدم دیتا را از منابع دیگری دریافت کنم و بعد به صورت محدود شافل کنم). بعد از آن بر روی این دیتا تغییرات را اعمال می‌کنم. دیتای تست را هم بر همین اساس و به صورت رندوم و با فرمت خواسته شده درآوردم. در مرحله بعد دیتا را آماده Fine-Tune کردن کردم و یک پرامپت چندخطه تشکیل دادم و بد با استفاده از کلاس Dataset در transformers به صورت یک دیتاست قابل خواندن برای این ماژول درآوردم. در نهایت دیتای اصلی را با استفاده از این تابع تغییر دادم و داده‌های آموزش را تشکیل دادم. در مرحله بعد تنظیمات کوانتیزاسیون را برای کاهش حجم مدل انجام دادم و مرحله بعد هم تنظیمات LoRA را برای Fine-Tuning درست کردم. در بخش بعد کارهای آموزش مدل را با استفاده از مدل Llama 3.2 با یک میلیارد پارامتر انجام دادم و مدل را برای آموزش با بیت‌های کم با استفاده از تابع prepare_model_for_kbit_training آماده کردم. همچنین با استفاده از get_peft_model تنظیمات LoRA را اعمال کردم. در آخرین مرحله هم یک تابع برای ارزیابی مدل زدم که سوال را استخراج می‌کند و به عنوان پرامپت به مدل می‌دهد و جواب را از مدل می‌گیرد اگر جواب داده شده درست باشد و بین عدد 1 تا 4 باشد جواب را برداشت می‌کنم در غیر اینصورت به عنوان غلط خروجی می‌دهم. همچنین در مرحله بعد جواب درست سوال و جواب مدل را با هم مقایسه می‌کنم و در صورت درست بودن خروجی درست می‌شود. در آخرین مرحله هم خطاها را بررسی می‌کنم و چک می‌کنم دلیل وجود خطاها چه بوده و هر کدام به چه کلمات کلیدی‌ای در سوال داشتند.

برای این مسئله نتیجه من در هر دو حالت 22 درصد بود که خب نتیجه قابل قبولی نیست و به نظر می‌رسد ایراد از نحوه برداشت دیتای من برای مسئله باشد چرا که تعریف مدل و پارامترها تا حد خوبی درست هستند و حداقل با چند بار آموزش دیدن باید نتیجه مدل متفاوت‌تر از حالت عادی و بیس باشد که اینگونه نیست.