این کد به پیاده‌سازی یک مدل یادگیری عمیق برای تبدیل تاریخ‌های غیررسمی به فرمت‌های رسمی پرداخته و با استفاده از روش‌های نوین NLP دقت بالایی در تولید خروجی‌های متنی دارد.

**1: وارد کردن کتابخانه‌ها**

* **pandas**: برای مدیریت و تحلیل داده‌های جدولی استفاده می‌شود.
* **sklearn.model\_selection**: برای تقسیم داده‌ها به مجموعه‌های آموزشی و تستی به کار می‌رود.
* **transformers**: این کتابخانه برای مدل‌های پیشرفته پردازش زبان طبیعی (NLP) به کار می‌رود و ما از مدل T5 استفاده می‌کنیم.
* **torch**: برای انجام محاسبات tensor و یادگیری عمیق استفاده می‌شود.

**.2 تولید مجموعه داده‌ها**

**2.1 : شناسایی فرمت‌های تاریخ**

* در این بخش، تاریخ‌ها به دو شکل غیررسمی و رسمی تعریف می‌شوند که پایه‌ای برای آموزش مدل هستند.

**:2.2 ایجاد DataFrame**

* با استفاده از دیکشنری شامل تاریخ‌های غیررسمی و رسمی، یک DataFrame ایجاد می‌شود که امکان پردازش داده‌ها را فراهم می‌کند.

**:2.3 ذخیره و نمایش DataFrame**

* DataFrame به فرمت CSV ذخیره می‌شود تا در مراحل بعدی از آن استفاده شود و همچنین برای بررسی صحت داده‌ها چاپ می‌شود.

**.3 آماده‌سازی داده‌ها**

**:3.1 بارگذاری مجموعه داده**

* داده‌های ذخیره‌شده در فایل CSV بارگذاری می‌شوند.

**:3.2 تقسیم داده‌ها**

* داده‌ها به دو بخش آموزشی (80%) و تستی (20%) تقسیم می‌شوند تا مدل بر اساس داده‌های آموزشی یاد بگیرد و سپس بر روی داده‌های تستی ارزیابی شود.

**:3.3 بارگذاری توکنایزر و مدل T5**

* توکنایزر برای تبدیل متون به توکن‌ها و مدل T5 برای تولید خروجی متنی بارگذاری می‌شود.

**3.4 :توکنایز کردن داده‌ها**

* داده‌های تاریخ به فرمت قابل فهم برای مدل تبدیل می‌شوند و با استفاده از truncation و padding، اطمینان حاصل می‌شود که طول توکن‌ها یکسان باشد.

**.4 آموزش مدل**

**:4.1 ایجاد DataLoader**

* با استفاده از ورودی‌ها و برچسب‌ها، مجموعه داده‌ای ایجاد می‌شود و DataLoader برای بارگذاری داده‌ها به صورت بچ‌های کوچک طراحی می‌شود.

**:4.2 تعریف بهینه‌ساز و حالت آموزش**

* AdamW به عنوان بهینه‌ساز انتخاب می‌شود و مدل به حالت آموزش تنظیم می‌شود.

**:4.3 آموزش مدل**

* حلقه‌ای برای آموزش مدل ایجاد می‌شود که در آن ورودی‌ها به مدل داده می‌شوند و بر اساس میزان خطا، وزن‌ها به‌روزرسانی می‌شوند.

**5 .ارزیابی مدل**

**:5.1 تنظیم مدل در حالت ارزیابی**

* مدل برای ارزیابی بر روی داده‌های تست تنظیم می‌شود.

**:5.2 تولید پیش‌بینی‌ها**

* خروجی‌های پیش‌بینی شده برای ورودی‌های تست تولید می‌شود.

**:5.3 رمزگشایی خروجی‌ها**

* توکن‌های تولید شده به متن قابل خواندن تبدیل می‌شوند.

**.6 اصلاح فرمت تاریخ پیش‌بینی شده**

* تابعی برای اصلاح پیش‌بینی‌های تاریخ ایجاد می‌شود که با استخراج سال و روز، فرمت تاریخ نهایی را تنظیم می‌کند.

**.7 نمایش نتایج**

* نتایج شامل تاریخ‌های غیررسمی، رسمی و پیش‌بینی‌های مدل چاپ می‌شود تا دقت مدل مورد بررسی قرار گیرد.

.