

## تمرین هشتم

---

**هدف:** آشنایی با شبکه ترنسفورمرز.

**کد:** کد این فعالیت را به زبان پایتون و با استفاده از چارچوب‌های کاری تنسورفلو یا پای‌تورچ بنویسید.

**گزارش:** ملاک اصلی انجام فعالیت گزارش آن است و ارسال کد بدون گزارش فاقد ارزش است. برای این فعالیت یک فایل گزارش در قالب pdf تهیه کنید و در آن برای هر سوال، تصاویر ورودی، تصاویر خروجی و توضیحات مربوط به آن را ذکر کنید. سعی کنید توضیحات کامل و جامعی تهیه کنید.

**تذکر ۱:** مطابق قوانین دانشگاه هر نوع کپی برداری و اشتراک کار دانشجویان غیر مجاز بوده و شدیداً برخورد خواهد شد. استفاده از کدها و توضیحات اینترنت به منظور یادگیری بلامانع است، اما کپی کردن غیرمجاز است.

**تذکر ۲:** مجموعه‌های داده مورد استفاده را به جز در مواردی که صریحاً در صورت سوال ذکر شده باشد، حتماً قبل از استفاده بصورت تصادفی به سه بخش آموزش (۷۰ درصد داده‌ها)، آزمون (۲۰ درصد داده‌ها) و اعتبارسنجی (۱۰ درصد داده‌ها) تقسیم نمایید.

**تذکر ۳:** مدل‌های تخمینگر را بر اساس معیار میانگین مربعات خطا و مدل‌های دسته‌بند را براساس دقت و ماتریس درهم‌ریختگی ارزیابی نمایید.

**راهنمایی:** در صورت نیاز میتوانید سوالات خود را در خصوص پروژه از تدریس‌یار درس، از طریق ایمیل زیر بپرسید.

E-mail: ann.ceit.aut@gmail.com

**ارسال:** فایل‌های کد و گزارش خود را در قالب یک فایل فشرده با فرمت StudentID\_HW۰۸.zip تا تاریخ ۱۴۰۱/۰۴/۲۴ ارسال نمایید. شایان ذکر است هر روز تاخیر باعث کسر ۱۰٪ نمره خواهد شد.

**توجه:** از آن‌جا که آموزش مدل ترنسفورمر، زمان و منابع زیادی لازم دارد، می‌توانید به جای آموزش مدل روی تمام داده‌ها، فقط از بخش کوچکی از داده‌ها به انتخاب خود استفاده کنید. با این کار، زمان اجرای آزمایشات کاهش می‌یابد و می‌توانید بیشتر به تحلیل عملکرد مدل بپردازید.

در این تمرین قصد داریم با استفاده از شبکه ترنسفورمرز یک مدل زبانی برای ترجمه ماشینی از زبان پرتغالی به انگلیسی بسازیم و عملکرد این شبکه را در ترجمه ماشینی مورد بررسی قرار دهیم. مجموعه‌داده

مورد استفاده در این تمرین را می‌توانید از طریق مجموعه‌داده‌های استاندارد موجود در تنسورفلو با نام «ted\_hrlr\_translate/pt\_to\_en» بارگذاری نمایید<sup>۱</sup>. این مجموعه‌داده شامل حدود ۵۰ هزار جمله آموزشی، ۱۱۰۰ جمله اعتبارسنجی و ۲۰۰۰ جمله تست می‌باشد.

برای انجام این تمرین به سوالات زیر پاسخ دهید.

---

<sup>۱</sup> tfds.load('ted\_hrlr\_translate/pt\_to\_en', with\_info=True, as\_supervised = True)

۱- مجموعه داده معرفی شده را بارگذاری نموده و کلمات موجود در جملات را با استفاده از یک توکنایزر، تبدیل به توکن نمایید. پیش پردازش های لازم را انجام داده و در فایل گزارش ذکر نمایید. (بعنوان مثال، می توانید جملات خیلی طولانی را کوتاه نمایید).

۲- در مکانیزم توجه، ورودی ها بعنوان یک مجموعه بدون ترتیب در نظر گرفته می شوند. اما در این مساله ورودی ها دنباله ای از کلمات هستند که ترتیب آن ها مهم است. برای حل کردن این مشکل می توان از یک تعبیه مکانی<sup>۲</sup> استفاده کرد. تعبیه مکانی را مطابق با روابط زیر محاسبه نمایید.

$$PE_{(pos,2i)} = \sin(pos/10000^{2i/d_{model}})$$
$$PE_{(pos,2i+1)} = \cos(pos/10000^{2i/d_{model}})$$

- ۳- ساختار شبکه ترنسفورمر را پیاده نموده و آن را روی مجموعه داده آموزش دهید. به سوالات زیر پاسخ دهید.
- بهینه ساز مناسب برای آموزش این مدل چیست؟ در صورتی که از بهینه سازی غیر از موارد استاندارد موجود در تنسورفلو استفاده می نمایید، ساختار کامل بهینه ساز را توضیح داده و آن را پیاده سازی نمایید.
  - پس از آموزش کامل مدل، نقشه حرارتی توجه روی هر کدام از head ها را برای یک جمله ورودی رسم نموده و آن را تحلیل نمایید.
  - ۲۰ نمونه از جملات موجود در مجموعه داده تست را به مدل داده و خروجی مدل و خروجی مطلوب را در فایل گزارش ذکر نمایید.

موفق باشید

---

<sup>۲</sup> Positional embedding