

به نام خدا

## تمرین سوم درس آنالیز داده

تاریخ تحویل: --

### مجموعه داده

Radiology Object in COntext Version 2 (ROCOv2). یک مجموعه داده چندرسانه‌ای متشکل از تصاویر رادیولوژی و مفاهیم و عناوین پزشکی مرتبط استخراج شده از زیرمجموعه دسترسی آزاد PubMed است. مفاهیم برای تشخیص بالینی، آناتومی (اشعه ایکس) و جهت (اشعه ایکس) به صورت دستی گردآوری و همچنین توسط یک متخصص رادیولوژی ارزیابی شده است. برخلاف MIMIC-CXR، ROCov2 شامل هفت روش بالینی مختلف است.

این مجموعه داده از آدرس زیر قابل دسترسی است:

<https://zenodo.org/records/10821435>

### بخش اول

موضوع: آشنایی با Masked Autoencoder و آموزش آن برای حوزه‌ی تصاویر پزشکی

### شرح تمرین

۱. تصاویر مربوط به قسمت Train و Test را از مجموعه داده‌ی ROCov2 بارگیری کنید.
۲. یک مدل از پیش آموزش دیده‌ی ViTMAE از کتابخانه‌ی Transformers (معرفی شده توسط huggingface) را load کرده و عملکرد آن را در بازسازی 5 نمونه از تصاویر موجود در Test با اعمال ۷۵ درصد Mask بررسی کنید.

راهنمایی:

[https://github.com/NielsRogge/Transformers-Tutorials/blob/master/ViTMAE/ViT\\_MAE\\_visualization\\_demo.ipynb](https://github.com/NielsRogge/Transformers-Tutorials/blob/master/ViTMAE/ViT_MAE_visualization_demo.ipynb)

۳. با استفاده از مجموعه‌ی Train و کتابخانه‌ی Transformers مدل ViTMAE برای حوزه‌ی تصاویر پزشکی finetune کنید. (برای حداقل ۵ epoch)

راهنمایی: برای غلبه بر محدودیت زمانی colab، مدل را در google drive ذخیره و دوباره load کنید.

۴. مدل finetune شده‌ی ViTMAE را load کرده و عملکرد آن را در بازسازی 5 نمونه از تصاویر MASK شده (۷۵ درصد) موجود در Test بررسی کنید.

۵. (اختیاری) به دلخواه یک دادگان جدید طبقه‌بندی تصاویر پزشکی انتخاب کنید.

۶. (اختیاری) با استفاده از انکودر ViTMAE ی finetune شده و آموزش مجدد برای دادگان جدید، نتایج زیر را گزارش دهید.

- Overall Accuracy

- Precision

- Recall

- F1-Score

- Confusion Matrix

## بخش دوم

موضوع: finetune کردن یک مدل Multi modal برای Medical Image Captioning با استفاده از LoRA

۱. با مراجعه به spreadsheet زیر مدلی که قصد finetune کردن آن را دارید، در مقابل اطلاعات گروه خود بنویسید. (هر گروه فقط یک مدل با وزن‌های Pretrain شده‌ی خاص را می‌تواند انتخاب کند. یعنی، امکان انتخاب مدل یکسان ولی با اندازه و وزن‌های متفاوت وجود دارد).

<https://docs.google.com/spreadsheets/d/1H0Z2vtqy3zTaV6LktORwf-T-SSBqUS65Qm3rAoeCVTA/edit?usp=sharing>

۲. مجموعه‌های Train، Test و کپشن‌های آن‌ها را از مخزن دادگان ROCov2 دانلود کنید.

۳. مدل پیش آموزش دیده را لود کرده و با افزودن LoRA از کتابخانه‌ی PEFT به آن مدلی جدید بسازید.

راهنمایی:

[https://github.com/huggingface/notebooks/blob/main/peft/Fine\\_tune\\_Blip2\\_on\\_an\\_image\\_captioning\\_dataset\\_PEFT.ipynb](https://github.com/huggingface/notebooks/blob/main/peft/Fine_tune_Blip2_on_an_image_captioning_dataset_PEFT.ipynb)

۴. با استفاده از تصاویر و کپشن‌های مجموعه‌ی Train، مدل جدید را finetune کنید.

راهنمایی: برای غلبه بر محدودیت زمانی colab، مدل را در google drive ذخیره و دوباره load کنید.

۵. با استفاده از مدل finetune شده و تصاویر و کپشن‌های مجموعه‌ی Test، معیارهای ارزیابی زیر را برای مدل خود به دست آورید.

- CIDEr
- BLEU
- METEOR
- ROUGE

## نکات تکمیلی

قبل از انجام تمرین:

- حتما لینک‌های فرستاده شده به عنوان راهنمایی را مطالعه فرمایید.
- حتما آموزش‌های مربوط به کتابخانه‌های Transformers و PEFT را مطالعه کنید.
- در مورد مدل‌های multi modal کتابخانه‌ی Transformers و امکان استفاده از آن‌ها برای image captioning اطلاعات کسب کنید.

در بخش دوم:

- برای افزایش سرعت آموزش، بهتر است از مدل‌های با تعداد پارامتر کمتر از ۳ میلیارد استفاده کنید.
- در صورتی که از نظر زمان اجرایی colab به مشکل برخوردید، می‌توانید نمونه‌های مجموعه داده‌ی Train را کاهش دهید.
- در صورتی که از نظر زمان اجرایی colab در محاسبه‌ی metric ها به مشکل برخوردید، حتما در گروه اعلام کنید.

## ضوابط ارسال تمرین

۱. مشاهدات خود و نتایج خواسته شده در هر قسمت را در قالب یک گزارش کامل ارسال کنید. گزارش ارسالی مبنای نمره دهی و تصحیح خواهد بود.
۲. نیازی به افزودن و توضیح کد در متن گزارش نیست، اما کدهای خود را چه به صورت jupyter notebook یا فایل python همراه با گزارش ارسال کنید.
۳. در صورت مشاهده‌ی مشابهت غیر قابل قبول بین گزارش و کد ارسالی با کارهای دیگر، حداقل نمره در نظر گرفته خواهد شد. لطفا فقط کار خود را ارائه دهید.

۴. تمامی فایل های تمرین را به صورت یک فایل archive با نام HW3\_(FullName)\_(StudentID).zip بارگذاری کنید.

**موفق باشید.**