

## سوال ورکشاپ VQA:

ابتدا در نوت بوک مربوط به ورکشاپ، تعداد کلاس ها را برابر با 5 قرار دهید. به جای استفاده از Adam به عنوان بهینه ساز، از AdamW استفاده کنید. به جای استفاده از resnet18 از resnet50 استفاده کنید (احتمالا نیاز خواهد بود تا ابعاد متغیر hidden\_dim را تغییر دهید). سپس مطابق لینک زیر، پارامترهای مدل image\_feature\_extractor را freeze کنید:

[https://jimmy-shen.medium.com/pytorch-freeze-part-of-the-layers-4554105e03a6#:~:text=%3E%3E%3E%20for%20name%2C%20parameters\(\)%3A%0A...%20%20%20%20%20para.requires\\_grad%20%3D%20False%0A...%20%20%20%20%20print\(%22%2D%22\\*20\)%0A...%20%20%20%20%20print\(f%22name%3A%20%7Bname%7D%22\)%0A...%20%20%20%20%20print\(%22values%3A%20%22\)%0A...%20%20%20%20%20print\(para\)](https://jimmy-shen.medium.com/pytorch-freeze-part-of-the-layers-4554105e03a6#:~:text=%3E%3E%3E%20for%20name%2C%20parameters()%3A%0A...%20%20%20%20%20para.requires_grad%20%3D%20False%0A...%20%20%20%20%20print(%22%2D%22*20)%0A...%20%20%20%20%20print(f%22name%3A%20%7Bname%7D%22)%0A...%20%20%20%20%20print(%22values%3A%20%22)%0A...%20%20%20%20%20print(para))

سپس مطابق لینک زیر، یک scheduler نوشته (CyclicLR - exp\_range) و برای آن پارامترهای زیر را تعیین کنید:

<https://www.kaggle.com/code/isbhargav/guide-to-pytorch-learning-rate-scheduling>

```
base_lr=0.001, max_lr=0.1, step_size_up=71, mode="exp_range",  
cycle_momentum=False
```

فقط به این نکته توجه کنید که **بعد از انجام ارزیابی در هر ایستگاه**، learning rate را به روز رسانی کنید. زیرا در غیر اینصورت، ارزیابی عادلانه نخواهد بود.

## سوال تمرین 5 (question answering):

در فایل Q9.ipynb قسمتهای خواسته شده را با مدل roberta-base تکمیل کنید.