

# بسم الله الرحمن الرحيم

دانشگاه علم و صنعت ایران

بهار ۱۴۰۰

تمرین سری دهم

تحويل: دوشنبه ۱۰ خرداد

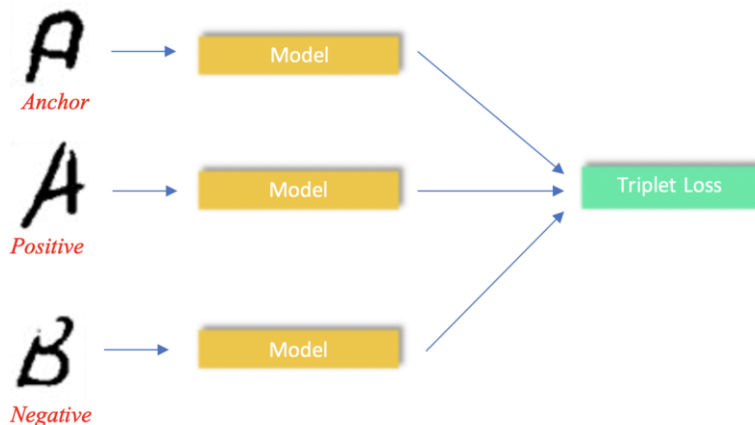
یادگیری عمیق

۱. دو روش skip-gram و CBOW در word embedding جهت انجام tokenization را مقایسه کنید و مزایا و معایب آنها را بیان نمایید.

مطالعه بخش هایی از مقاله زیر جهت پاسخ به این پرسش توصیه میشود.

Camacho-Collados, Jose, and Mohammad Taher Pilehvar. "[From word to sense embeddings: A survey on vector representations of meaning](#)." *Journal of Artificial Intelligence Research* 63 (2018): 743-788.

۲. شکل زیر را در نظر بگیرید. فرض کنید در این مثال  $d(A, P) = 0.4$  و  $d(A, N) = 0.5$  باشد و  $\alpha = 0.2$  انتخاب شده باشد (فرض کنید در محاسبه  $d$  مطابق با اسلاید از نرم ۲ استفاده می شود).  
الف) مقدار تابع ضرر Triplet برای این نمونه را محاسبه کنید.  
ب) گرادیان تابع ضرر نسبت به embeddingهای مربوط به تصاویر  $A$ ،  $P$  و  $N$  را محاسبه کنید. منظور از embedding در این قسمت خروجی مدل به ازای تصویر ورودی است که فرض کنید یک بردار ۲۵۶ بعدی است (به ترتیب با نمادهای  $e_A$ ،  $e_P$  و  $e_N$  نشان داده می شوند) و هر کدام از این گرادیانها نیز یک بردار ۲۵۶ بعدی خواهند بود (آنها را با نمادهای  $g_A$ ،  $g_P$  و  $g_N$  نشان دهید) که توابعی از  $e_A$ ،  $e_P$  و  $e_N$  خواهند بود.  
پ) اگر  $d(A, N) = 1.5$  بود، نتیجه قسمت های (الف) و (ب) را مجدداً محاسبه کنید.



۳. این نوت بوک شامل یک روش end to end برای پیاده سازی image caption generation بر روی دیتاست MS-coco مشابه مقاله زیر است:

Xu, Kelvin, et al. "[Show, attend and tell: Neural image caption generation with visual attention](#)." *International conference on machine learning*. PMLR, 2015.

مقاله را مطالعه نمایید و کد ارسالی را اجرا کنید و گزارش کاملی شامل موارد زیر ارسال نمایید. (طبق مقاله و تحلیل کد)

- مجموعه داده
- پیش پردازش
- استخراج ویژگی (استفاده از inception v3)

- Tokenization
- مدل و معماری شبکه
- بررسی مکانیزم attention
- نتایج

## نکات تکمیلی

- (۱) لطفاً پاسخ سوالات (تئوری و توضیحات پیاده‌سازی) را به طور گویا و به زبان فارسی و در صورت امکان تایپ همراه با سورس کدهای نوشته شده، در یک فایل فشرده شده به شکل HW10\_YourStudentID.zip قرار داده و بارگذاری نمایید.
- (۲) منابع استفاده شده را به طور دقیق ذکر کنید.
- (۳) برای سهولت در پیاده‌سازی‌ها و منابع بیشتر، زبان پایتون پیشنهاد می‌شود. لطفاً کدهای مربوطه را به طور جداگانه در فرمت py یا ipynb ارسال نمایید.
- (۴) ارزیابی تمرین‌ها براساس صحیح بودن راه حل‌ها، گزارش مناسب، بهینه بودن کدها و کپی نبودن می‌باشد.
- (۵) در مجموع تمام تمرین‌ها، تنها ۷۲ ساعت تاخیر در ارسال پاسخ‌ها مجاز است اما پس از آن به صورت خطی از نمره کسر خواهد شد (معادل با روزی ۵۰ درصد).
- (۶) در رابطه با پرسش و پاسخ در رابطه با تمرین‌ها می‌توانید در گروه مربوطه مطرح کنید.

موفق باشید.