

| | |
|-----------------------------------|--|
| <p>تمرین سری ۶</p> <p>درس NLP</p> | <p>نام مدرس: دکتر مرضیه داودآبادی فراهانی</p> <p>دستیاران آموزشی: امیرمحمد آزادی، بکتاش انصاری، سینا زمانی</p> |
|-----------------------------------|--|

سوالات تئوری

۱. ✓ دو سناریوی آموزش مدلی مانند BERT را در نظر بگیرید که در اولی وزن‌های اولیه‌ی آموزش برابر با وزن‌های یک مدل BERT از قبل آموزش دیده شده (Pretrain) بر روی مجموعه تسک‌های مختلف می‌باشد. و در دومی وزن‌های اولیه مقادیری تصادفی می‌باشند. توضیح دهید هر کدام از مدل‌ها در فرآیند آموزش به طور تقریبی چگونه عمل خواهند کرد و همین‌طور انتظار داریم عملکرد مدل‌ها پس از فرآیند آموزش چگونه باشد؟
۲. ✓ یکی از چالش‌های قابل توجه در فرآیند آموزش یک مدل شبکه عصبی بخصوص در فرآیند Fine-Tuning، Catastrophic Forgetting می‌باشد. این چالش را به طور کامل توضیح دهید و روشی ارائه دهید که بتوانیم این چالش را کاهش دهیم و یا برطرف کنیم. (منابعی که در آن‌ها جست و جو کردید را ارائه دهید.)
۳. ✓ در رابطه با Transfer Learning و تفاوت‌های آن با Fine-Tuning تحقیق کنید و توضیح دهید در چه شرایطی از هر کدام استفاده میشود؟
۴. ✓ راجع به تاثیرات روش‌های مختلف masking و همین‌طور تعیین میزان توکن‌های قابل mask بر روی فرآیند آموزش و عملکرد مدل‌های MLMS توضیح دهید. (روش رندوم و روش مبتنی بر part of speech)

۵. ✓ عملکرد معماری‌های CLM, MLM, Seq2Seq را با یکدیگر مقایسه کنید و مزایا و معایب هر کدام را بیان کنید. و از هر کدام نمونه‌هایی را مثال بزنید.

۶. ✓ با توجه به عملکرد مدل‌های MLM (پیشبینی توکن‌های mask شده)، چگونه میتوان از آن‌ها برای تولید یک دنباله‌ای از متن استفاده کرد؟

سوالات عملی

۷. ✓ در نوت بوک q7.ipynb میخواهیم معماری یک مدل MLM را پیاده سازی کنیم. آن را تکمیل کنید. (سوالات تشریحی درون نوت بوک را در پی‌دی‌اف پاسخ سوالات تئوری قرار دهید)