

به نام خدا

تمرین دوم درس فهم زبان، پاییز ۱۴۰۱

دانشکده مهندسی کامپیوتر، دانشگاه صنعتی امیرکبیر

در این تمرین به مسئله **پرکردن شکاف^۱ و تشخیص قصد^۲** پرداخته خواهد شد. برای این تمرین از مجموعه داده ضمیمه شده استفاده خواهیم کرد که همان مجموعه داده پرکردن شکاف و تشخیص قصد چند زبانه فیسبوک است. در این مجموعه داده به ازای هر یک از جملات یک برچسب قصد و به ازای هر یک از کلمات یک جمله، یک برچسب شکاف اختصاص داده شده است که تعداد زیادی از آنها برچسب تهی دارند. این مجموعه به سه بخش مجموعه آموزش (train)، توسعه (development) و آزمون (test) تقسیم شده است. از مجموعه داده آموزش و توسعه برای دستیابی به بهترین مدل استفاده کرده و از مجموعه داده آزمون برای سنجش و ارزیابی مدل استفاده کنید. جزئیات دادگان را می‌توانید در مقاله خود فیسبوک در آدرس زیر ببینید.

<https://arxiv.org/abs/1810.13327>

در این جا شما باید مدلی طراحی کنید که به ازای هر یک از جملات مجموعه داده بتواند برچسب قصد و برچسب‌های شکاف‌ها را تشخیص دهد. در این مدل باید از روش‌های مبتنی بر مشتقات شبکه‌های عصبی ترنسفورمری استفاده شود و شما مجاز هستید تا از تمامی شبکه‌های ترنسفورمری شناخته شده استفاده کنید. از آن جایی که برای مدل‌های مبتنی بر ترنسفورمر باید از Tokenizerهای مبتنی بر زیرکلمه^۳ استفاده شود، شما باید به نحوه مقتضی توکن‌های مورد نیاز برای جمله را ایجاد کرده و برچسب‌های شکاف‌ها را نیز متناسب با آن‌ها هماهنگ کنید.

برای این تمرین باید حداقل دو مدل مختلف را بررسی و نتایج آن‌ها را با هم مقایسه کنید. یکی از این روش‌های می‌تواند بر مبنای ترنسفورمر نباشد. همچنین دیگر مقایسه‌ها از نظر تعداد پارامتر و زمان‌های آموزش و تست نیز باید انجام شود. با توجه به مطالبی که در درس یادگرفته‌اید توقع است که نوآوری‌هایی در پیاده‌سازی‌ها داشته باشید که باعث بهبود دقت شوند. این نوآوری‌ها می‌تواند به عنوان بخش پیشنهاد ایده جدید هم مورد استفاده قرار گیرد در صورتی که منطق خوبی داشته باشد و بتوان از آن دفاع کرد.

برای ارزیابی مدل می‌توانید از قطعه کد زیر که به زبان perl نوشته شده است استفاده نمایید. این قطعه کد مبتنی بر استاندارد Conll معیارهای F1، Precision و Recall را محاسبه می‌کند.

https://www.comp.nus.edu.sg/~kanmy/courses/practicalNLP_2008/packages/conlleval.pl

برای استفاده از این قطعه کد در زبان پایتون از فایل eval.py موجود در فایل‌های ضمیمه استفاده کنید.

¹Slot Filling

²Intent detection

³ Sub-word

بعد از انجام تمرین باید یک گزارش مناسب (شامل بخش‌های مورد نیاز و با توضیحات و نتیجه‌گیری مناسب) برای تمرین بنویسید. تمرین فاقد گزارش بدون بررسی نمره صفر خواهد گرفت. گزارش ۳۰ درصد نمره تمرین را دارد. همچنین تمرین باید به طور کامل کد پایتون به صورت ماژولار باشد. یعنی اینکه نوتبوک و یا تبدیل نوتبوک به پایتون به صورت اتوماتیک قابل قبول نیست و بدون بررسی صفر داده می‌شود. ۲۰ درصد نمره به نتایج بدست آمده اختصاص دارد و ۵۰ درصد تمرین هم به پیاده‌سازی مناسب و کدنویسی تمیز و ماژولار داده خواهد شد. ددلاین تمرین سه هفته است و تا دو روز با کسر نمره یک درصد به ازای هر ساعت می‌توانید با تاخیر ارسال کنید. بعد از این زمان دیگر امکان ارسال پاسخ تمرین وجود ندارد.

برای اینکه ترغیب شوید به انجام هرچه بهتر تمرین و انتخاب بهترین استراتژی‌ها، به نفرات اول تا سوم بر اساس نرخ خطای واج به ترتیب ۱۵، ۱۰ و ۵ درصد نمره اضافه داده خواهد شد. البته باید نتایج از یک حدی بهتر باشند (یعنی اینکه اگر همه پیاده‌سازی‌ها مناسب نباشد، دلیلی برای دادن نمره اضافه نیست).

سوالات خود در رابطه با تمرین را می‌توانید در گروه تلگرام بپرسید.

موفق باشید

زینلی