



مینی پروژه پردازش زبان طبیعی

توجه:

- کد باید فقط در زبان پایتون باشد.
 - استفاده از کتابخانه‌های از پیش آماده مجاز است. در گزارش کتابخانه‌ها ذکر شوند.
 - لطفاً علاوه بر ضمیمه کردن کد، نتایج را تحلیل و در فایل گزارش خود ضمیمه کنید.
 - فایل گزارش به فرمت pdf و به زبان فارسی باشد.
-
- در این مینی پروژه با استفاده از تنها یک شبکه عصبی مسئله POS و NER را با حل می کنید.
 - دادگان به سه دسته آموزشی، اعتبارسنجی و تست تقسیم شده و در لینک قرار دارند و با استفاده از پایتون pickle شده اند. هر بخش داده دارای ۳ فیلد `tokens`، `ner_tags` و `pos_tags` است، اما دادگان تست برچسب ندارند.
 - مدلی بر پایه شبکه پیش آموزش دیده bert-base-cased برای پیش بینی POS و NER طراحی کنید (این مدل پیش آموزش دیده فقط نقش backbone را دارد، در نتیجه شما آزادی عمل در طراحی مدل کلی را دارید). در نظر داشته باشید که مدل طراحی شده باید به صورت پایان-به-پایان^۱ دو برچسب را پیش بینی کند (Multi-Tasking Learning).
 - مدل و روش های تان با استفاده از معیار های دقت، recall، precision و f1-score مورد ارزیابی قرار دهید و گزارش کنید (از میانگین micro استفاده کنید).
 - مدل بهینه تان را انتخاب کنید و بر روی دادگان تست پیشبینی انجام دهید. پیشبینی را با فرمت مشابه برچسب های داده آموزشی pickle کرده و با نام `test_pred.pickle` ضمیمه کنید.
 - گزارش کامل از پیش پردازش، طراحی مدل، نحوه آموزش و ... به همراه کد و پیش بینی های تان ضمیمه کنید.
 - در نظر داشته باشید که میزان دقت بر روی دادگان تست و گزارش شما در نمره نهایی پروژه تاثیر بالایی خواهد داشت.