به نام خدا

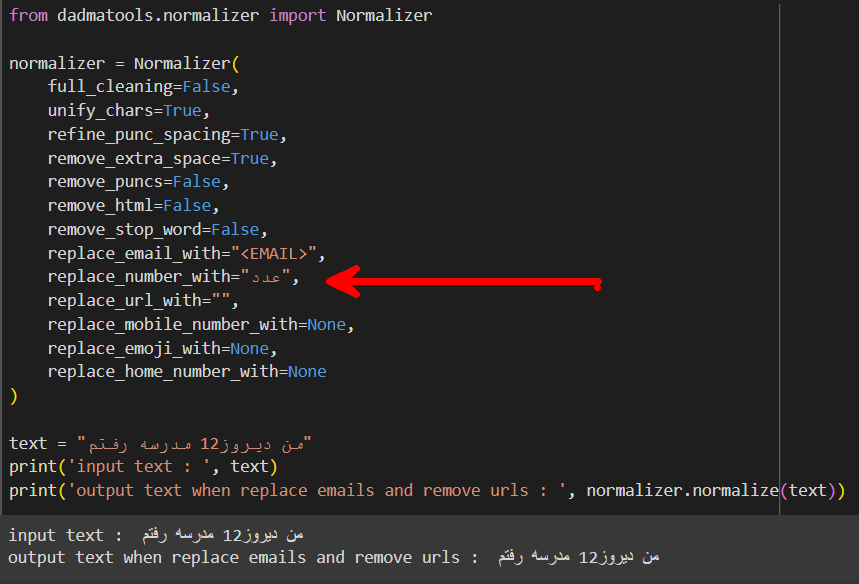
درس مبانی پردازش زبان و گفتار

کارگاه ابزار دادماتولز

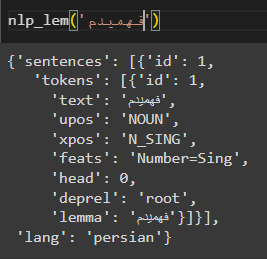
سینا علی‌نژاد

99521469

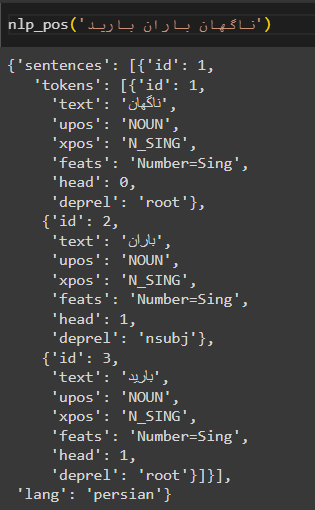
* برای ماژول normalizer مورد زیر را یافتم که نتوانسته عدد 12 را با کلمه‌ی "عدد" جایگزین کند.



* برای ماژول lemmatize فعل "فهمیدم" را دادم و خروجی نیز همان "فهمیدم" شد. در حالیکه باید فهمید/فهم میشد.



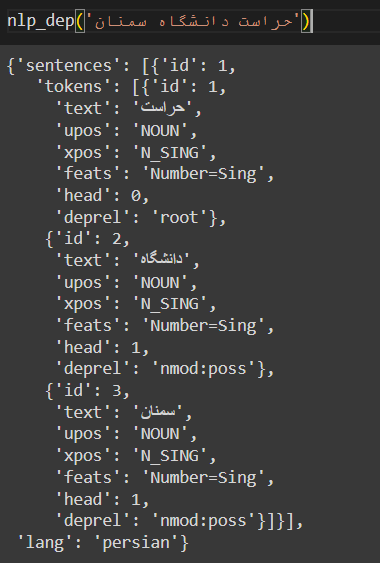
* برای عملیات part of speech tagging جلمه "ناگهان باران بارید" را دادم و برای کلمه‌ی ناگهان پیشبینی نادرست انجام شد. باید ADV بدهد اما Noun داد.



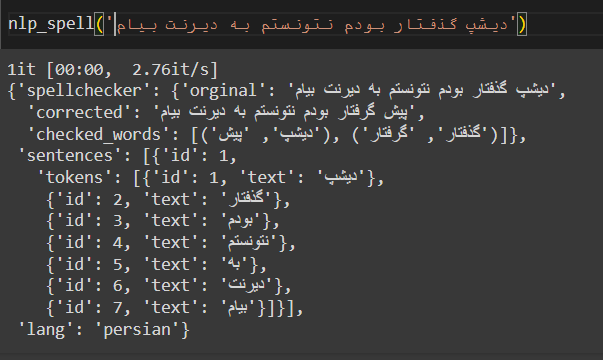
* برای عملیات named entity recognition در جمله ای که از عبارت "دکتر احمدی" استفاده کردم، نتوانست متوجه شود که دکتر احمدی با همدیگر یک entity است و فقط کلمه‌ی "احمدی" را به عنوان شخص شناخت. برای دیگر عناوین مثل پروفسور سمیعی نیز به همین صورت بود.



* برای گراف وابستگی، بهش جمله "حراست دانشگاه سمنان" رو دادم. اینجا باید بدین شکل عمل بکنه که کلمه "حراست" ریشه هست و کلمه "دانشگاه" وابسته به کلمه "حراست" هست و کلمه "سمنان" وابسته به کلمه "دانشگاه" هست. اما خروجی یک اشتباه دارد و آن هم روی کلمه "سمنان" است که آن را وابسته کلمه "حراست" می‌داند.



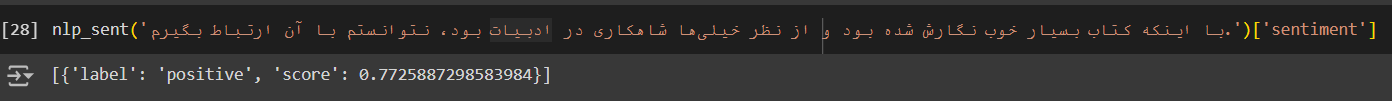
* در جمله زیر، نتوانسته متوجه شود "دیشپ" همان "دیشب" است. همچنین "دیرنت" همان "دیدنت" است.



* در جمله زیر، در دو جا نتوانسته تشخیص دهد کلمه نیاز به کسره دارد. یکی در عبارت "کلید در" روی کلمه "کلید" و دیگری در عبارت "حراست دانشگاه علم و صنعت" روی کلمه "دانشگاه"



* در جمله زیر، پس از بیان اینکه کتاب مذکور بسیار محبوب است و نگارشش خوب بوده، نظر خودم را گفتم که با آن ارتباط نگرفتم. پس باید احساس منفی داشته باشد ولی مدل احساس مثبت را با احتمال 77 درصد پیش بینی کرده است.



* در عبارت "خیلی خستم" نتوانسته است کلمه "خستم" را به شکل رسمی آن یعنی "خسته‌ام" یا "خسته هستم" تغییر دهد.

