

به نام خدا

تمرین دوم درس پردازش زبان های طبیعی

جواب خود را به همراه کد و گزارش کار شامل توضیحات کامل تحویل دهید.
(استفاده از Chat GPT بلامانع است).

سوال ۱: الگوریتم Viterbi را به طور کامل پیاده سازی کنید. سپس از یک corpus (از دیتابیس treebank نیز می‌توانید استفاده کنید) برای یادگیری احتمالات جدول A, B, π استفاده کنید. در نهایت مدل بدست آمده را بر روی دیتای تست ارزیابی کنید و نتایج دقت را گزارش کنید.

سوال ۲: فایل پیوست شده را بخوانید، تابعی بنویسید که فایل را دریافت کند و gramer آن را تحویل دهد. سپس به وسیله این گرامر درخت پارس هر کدام از جملات را بتواند در خروجی پرینت کند. برای این کار از کتابخانه‌های NLTK.CFG، NLTK.PARSE.ChartParser، NLTK.Pos_tag. می‌توانید استفاده کنید.

سوال ۳: از کتابخانه NLTK.Tag، hmm را به پروژه اضافه کنید. سپس مدل سوپروایزی بسازید که به وسیله دیتاهای استفاده شده در سوال ۱ پروسه یادگیری را انجام و ارزیابی کند. پس از آن نتایج حاصل از این ارزیابی را با نتایج سوال ۱ مقایسه و گزارش کنید.