

دانشکده مهندسی کامپیوتر

تمرین سری پنجم

درس پردازش زبان های طبیعی

پوریا رحیمی - 99521289

1403-1404

سوال اول) دو نوع از سیستم های Question Answering عبارتند از open-domain و close-domain در مورد هر کدام توضیح دهید و تفاوت آنها را بیان کنید.

سیستم‌های پاسخگویی به سوالات (Question Answering) به دو دسته کلی تقسیم می‌شوند: Open-domain و Closed-domain

1. **: Open-domain QA**این نوع سیستم‌ها سعی دارند به سوالات در مورد هر موضوعی پاسخ دهند. این سیستم‌ها باید دانش گسترده‌ای داشته باشند زیرا موضوعات ممکن است شامل هر چیزی از تاریخ و جغرافیا تا فرهنگ و علم باشند. این سیستم‌ها معمولاً از منابع متنوعی مانند اینترنت برای جمع‌آوری اطلاعات استفاده می‌کنند.
2. **Closed-domain QA** **:** این سیستم‌ها به سوالات در یک حوزه خاص یا محدود پاسخ می‌دهند. این سیستم‌ها معمولاً دقیق‌تر هستند زیرا دانش آن‌ها به یک حوزه خاص محدود شده است. به عنوان مثال، یک سیستم پاسخگویی به سوالات در حوزه پزشکی فقط به سوالات مربوط به پزشکی پاسخ می‌دهد.

تفاوت‌ها: تفاوت اصلی بین این دو نوع در حوزه دانش آن‌ها است. سیستم‌های Open-domain باید دانش گسترده‌ای داشته باشند و قادر به پاسخگویی به سوالات در موضوعات متنوع باشند، در حالی که سیستم‌های Closed-domain تنها به سوالات در یک حوزه خاص پاسخ می‌دهند. این باعث می‌شود سیستم‌های Closed-domain در حوزه خود دقیق‌تر باشند، اما آن‌ها نمی‌توانند به سوالات خارج از حوزه خود پاسخ دهند.

سوال دوم) ابتدا machine reading comprehension را تعریف کرده و سپس ارتباط آن را با Question Answering را توضیح دهید.

**Machine Reading Comprehension (MRC)**یکی از مسائل کلیدی در فهم زبان طبیعی است. [هدف از MRC آموختن به ماشین برای خواندن و درک متن داده شده و سپس پاسخ دادن به سوالات بر اساس آن است](https://paperswithcode.com/task/machine-reading-comprehension)  [MRC به ماشین کمک می کند تا متن غیرساختاریاده را بخواند و درک کند و پس از خواندن آن، بتواند به سوالات مرتبط با اطلاعات داده شده در فایل متنی پاسخ دهد](https://link.springer.com/chapter/10.1007/978-3-030-70542-8_20). [این هدف برای آموختن به ماشین‌ها برای درک متن مانند انسان وجود دارد و یک جهت چالش برانگیز جدید در هوش مصنوعی است](https://link.springer.com/chapter/10.1007/978-3-030-70542-8_20).

**Question Answering (QA)** یکی از برنامه های کاربردی اصلی MRC است. در QA، یک متن داده شده و سپس سوالاتی بر اساس آن متن پرسیده می شود. [سیستم باید بتواند به این سوالات پاسخ دهد](https://paperswithcode.com/task/machine-reading-comprehension). بنابراین، MRC و QA به طور مستقیم مرتبط هستند. [برای مثال، برخی از مدل های زبان پیش آموخته شده مانند Bert و BART، بر روی یک مجموعه داده بزرگ آموزش داده شده اند و بر روی داده های پایین دستی MRC تنظیم مجدد شده اند](https://paperswithcode.com/task/machine-reading-comprehension). [این مدل ها عملکرد بسیار خوبی در MRC نشان داده اند](https://paperswithcode.com/task/machine-reading-comprehension). بنابراین، می توان گفت که MRC یکی از روش های اصلی برای پیاده سازی سیستم های QA است.

سوال سوم) در حوزه QA ، دو نوع سوال factoid question و non-factoid question را توضیح دهید.

در حوزه پاسخ‌گویی به سوالات QA، سوالات به دو دسته کلی تقسیم می‌شوند :

Factoid Questions و Non-Factoid Questions

Factoid Questions :

این نوع سوالات به دنبال اطلاعات دقیق و مشخص هستند که معمولاً با یک جواب کوتاه و دقیق مانند یک نام، تاریخ، مکان یا عدد پاسخ داده می‌شوند. پاسخ به این سوالات معمولاً از طریق جستجو در یک پایگاه داده یا منابع اطلاعاتی به دست می‌آید. مثال‌هایی از این نوع سوالات عبارتند از :

اولین رئیس‌جمهور ایالات متحده کیست؟

پایتخت فرانسه کجاست؟

مساحت کشور چین چقدر است؟

Non-Factoid Questions :

این نوع سوالات به دنبال توضیحات، تحلیل‌ها، یا بحث‌های پیچیده‌تری هستند و نیاز به پاسخ‌های بلندتر و با جزئیات بیشتری دارند. پاسخ به این سوالات ممکن است شامل چندین جمله یا حتی چند پاراگراف باشد و ممکن است نیاز به ترکیب و تفسیر اطلاعات از منابع مختلف داشته باشد. مثال‌هایی از این نوع سوالات عبارتند از:

علت‌های اصلی جنگ جهانی دوم چه بود؟

تفاوت‌های کلیدی بین زبان‌های برنامه‌نویسی جاوا و پایتون چیست؟

سوالات non-factoid معمولاً چالش بیشتری برای سیستم‌های پاسخ‌گویی به سوالات دارند زیرا به تحلیل و درک عمیق‌تری از متن و موضوع نیاز دارند و ممکن است چندین دیدگاه یا جنبه مختلف را دربر گیرند.

سوال چهارم) مزایا و معایب transformer ها را نسبت به RNN ها بیان کنید.

مزایای Transformerها نسبت به RNNها :

1. مدیریت همزمان توالی‌ها :

Transformers : می‌توانند به طور همزمان تمام کلمات یک توالی را پردازش کنند، زیرا از مکانیزم Attention استفاده می‌کنند. این امکان پردازش موازی را فراهم می‌کند که سرعت آموزش و پردازش را به طور قابل توجهی افزایش می‌دهد.

RNNs : باید کلمات را به ترتیب زمانی پردازش کنند، به این معنی که کلمه بعدی نمی‌تواند قبل از پردازش کلمه قبلی پردازش شود. این ترتیب زمانی، سرعت پردازش را کاهش می‌دهد.

2. مدل‌سازی روابط بلندمدت :

Transformers به دلیل مکانیزم Attention، قادر به مدل‌سازی روابط طولانی مدت در توالی‌ها هستند، زیرا می‌توانند به راحتی به اطلاعات هر بخش از توالی دسترسی داشته باشند، بدون توجه به فاصله زمانی.

RNNs در مدل‌سازی روابط بلندمدت با مشکل مواجه می‌شوند، زیرا اطلاعات باید از طریق وضعیت‌های مخفی به جلو منتقل شوند که ممکن است منجر به مشکل فراموشی یا گرادیان محو شود.

3. کارایی محاسباتی:

Transformers به دلیل قابلیت پردازش موازی، کارایی بالاتری در استفاده از منابع سخت‌افزاری مدرن مانند GPU‌ها دارند.

RNNs به دلیل ماهیت ترتیبی پردازش، کارایی کمتری در بهره‌برداری از منابع محاسباتی موازی دارند.

معایب Transformerها نسبت به RNNها :

1. حساسیت به طول توالی:

Transformers ممکن است با افزایش طول توالی به مشکلاتی از قبیل نیاز به حافظه بیشتر و پیچیدگی محاسباتی بالاتر برخورد کنند. هر چه توالی طولانی‌تر باشد، محاسبات Attention نیز پیچیده‌تر و هزینه‌برتر می‌شود.

RNNs با اینکه پردازش ترتیبی دارند، طول توالی تأثیر مستقیم کمتری بر نیازهای حافظه و پیچیدگی محاسباتی آنها دارد.

2. نیاز به داده‌های زیاد:

Transformers برای آموزش بهینه نیاز به حجم زیادی از داده‌ها دارند تا بتوانند به خوبی تعمیم یابند.

RNNs معمولاً می‌توانند با داده‌های کمتر نسبت به Transformerها کار کنند، هر چند که این به قیمت کاهش عملکرد ممکن است تمام شود.

3. پیچیدگی معماری:

Transformers معماری پیچیده‌تری دارند و تنظیمات زیادی برای مکانیزم Attention و لایه‌های مختلف آنها وجود دارد که ممکن است تنظیم و بهینه‌سازی آنها را چالش‌برانگیز کند.

RNNs ساده‌تر هستند و پیاده‌سازی و تنظیم آنها ممکن است راحت‌تر باشد.

نتیجه‌گیری

در مجموع، Transformerها به دلیل توانایی پردازش موازی و کارایی بالا در مدل‌سازی روابط بلندمدت، انتخاب محبوب‌تری در بسیاری از کاربردهای NLP شده‌اند. با این حال، برای کاربردهایی که طول توالی‌ها کوتاه‌تر است یا منابع داده محدودتر هستند، RNNها همچنان می‌توانند گزینه مناسبی باشند.

سوال پنجم) در مورد positional encoding و اهمیت آن در مدل های مبتنی بر transformer توضیح دهید.

سوال ششم) در مورد تفاوت encoder-only ها، decoder-only ها و encoder-decoder ها در مدل های مبتنی بر transformer توضیح دهید.

سوال هفتم) یک دسته بندی دیگر برای سیستم های question answering عبارت است از extractive و abstractive. درمورد هر کدام و تفاوت آنها توضیح دهید.