

(۱)

پاسخگویی به سؤالات در دامنه بسته یا به عبارت دیگر Closed-Domain Question Answering (CDQA)، یک عبارت کلی است که برای پاسخگویی به سؤالاتی از یک حوزه خاص مانند حقوق، پزشکی، مهندسی و غیره استفاده می‌شود. مثلاً سوالی مثل "چه فرایندی برای انجام یک قرارداد لازم است؟" برای سیستم آموزش دیده برای حقوق.

در مقابل، پاسخگویی به سؤالات در دامنه باز یا به طور خلاصه Open-Domain Question Answering (OPQA) به معنای پاسخگویی به سؤالات در هر دامنه ای است. در OPQA، مدل آموزش دیده شده است تا بتواند به سؤالات مختلفی پاسخ دهد، بدون اینکه محدودیتی در حوزه سؤال اعمال شود. برای این کار، سیستم های OPQA ابتدا با استفاده از تکنیک های بازیابی اطلاعات، متن مرتبط با سؤال را پیدا کرده و سپس با استفاده از تکنیک های NLP بخش های مرتبط از آن را استخراج می کنند.

مثلاً سؤالاتی مثل: "چه زمانی استقلال ایران اعلام شد؟" یا "آیا هواپیماها در اسپانیا قرارداد دارند؟" تفاوت: هنگام آموزش یک سیستم CDQA، سؤالاتی که مدل می‌بیند، از یک مجموعه داده محدود در یک حوزه خاص است. اما در OPQA، مدل بر روی مجموعه داده های گسترده تری آموزش می‌بیند و محدودیت در حوزه سؤالات وجود ندارد.

(۲)

Machine Reading Comprehension یا MRC یا درک خواندن ماشین، وظیفه ای در پردازش زبان طبیعی (NLP) است که هدف آن ساختن سیستمی است که بتواند متن را برای پاسخ به سؤالات مربوط به آن درک کند. در این وظیفه، سیستم MRC دو ورودی دریافت می کند: سؤال و متن (context): متنی که ممکن است حاوی پاسخ سؤال باشد. هدف سیستم MRC این است که پاسخی را از متن استخراج کند یا بر اساس متن، پاسخی پیچیده تر ایجاد کند MRC به عنوان زیرمجموعه ای از پرسش و پاسخ (QA) محسوب می شود و بر نوع خاصی از پاسخگویی به سؤال تمرکز دارد.

تفاوت MRC با QA:

- QA: یک زمینه گسترده در NLP است که به طور کلی به سیستم هایی می پردازد که می توانند به سؤالات مطرح شده به زبان طبیعی پاسخ دهند.
- MRC: به طور خاص به وظایف QA اشاره دارد که در آن یک متن و سؤال مربوط به آن به سیستم داده می شود. هدف MRC این است که سیستم متن را به اندازه کافی بخواند و درک کند تا بتواند مستقیماً از اطلاعات موجود به سؤال پاسخ دهد.

بنابراین، همه وظایف MRC وظایف QA هستند، اما همه وظایف QA وظایف MRC نیستند. QA می تواند شامل منابع اطلاعاتی دیگری به غیر از متن مانند پایگاه های داده یا نمودارها باشد، در حالی که MRC به طور خاص بر توانایی خواندن و درک متن برای پاسخ به سؤالات تمرکز دارد.

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

سوالات حقیقتی (Factoid):

- این سوالات به دنبال پاسخ های قطعی و عینی هستند که به طور دقیق مشخص می شوند و معمولاً کوتاه و مختصر هستند و اغلب شامل نام، تاریخ، مکان یا اعداد هستند. آنها معمولاً شامل کلمات پرسشی مانند "چه"، "کی"، "کجا"، "چه کسی" یا "چند" هستند.
- مثال ها:

- پایتخت فرانسه کجا است؟ (پاسخ: پاریس)
- اولین فرود بر ماه چه سالی اتفاق افتاد؟ (پاسخ: 1969)
- بلندترین کوه جهان کدام است؟ (پاسخ: کوه اورست)

سوالات غیرحقیقتی (Non-Factoid):

- سوالات غیرحقیقتی فراتر از حقایق ساده رفته و به تفسیر، نظر یا استدلال می پردازند. پاسخ به این سوالات مفصل تر است و ممکن است نیاز به تجزیه و تحلیل، توضیح یا توجیه داشته باشد. آنها اغلب از عبارات چرا و چگونه استفاده می کنند و خواستار توضیح، توجیه یا فرضیه هستند.
- مثال ها:

- حل کردن یک مسئله ی ریاضی
- جواب دادن یک تست چهار گزینه ای
- چرا آسمان آبی است؟ (این پاسخ شامل توضیحاتی در مورد پراکندگی نور خواهد بود)
- موتور ماشین چگونه کار می کند؟ (این پاسخ نیازمند توضیح فرآیند احتراق است)

سیستم های پرسش و پاسخ حقیقتی (Factoid QA) برای پیدا کردن پاسخ های دقیق، بر روی تکنیک های بازیابی اطلاعات (IR) و پردازش زبان طبیعی (NLP) تکیه می کنند تا داده های خاصی را از متون و گاهی پایگاه های داده استخراج کنند. طراحی آنها معمولاً ساده تر است و می توانند برای پرسش های سرراست به دقت بالایی دست یابند.

در مقابل سیستم های پرسش و پاسخ غیرحقیقتی (Non-Factoid QA) به تکنیک های پیشرفته تر پردازش زبان طبیعی (NLP)، از جمله درک زمینه، استخراج اطلاعات مرتبط از متون طولانی تر و استدلال نیاز دارند. برای مدیریت پیچیدگی سوالات غیرحقیقتی، اغلب از مدل های مبتنی بر یادگیری عمیق مانند ترانسفورماتورها (BERT, GPT) استفاده می شود.

۴)

مزایا:

- مدل های Transformer می توانند به سادگی اطلاعات را به طور موازی در طول دنباله پردازش کنند. این موازی سازی باعث افزایش سرعت آموزش و پیش بینی می شود. در مقابل RNN ها برای مدل کردن وابستگی های زمانی در داده های ورودی مناسب تر هستند، زیرا اطلاعات ورودی به ترتیب به مدل معرفی می شوند.
- Transformer ها توانایی بهتری برای درک وابستگی های طولانی در داده های ورودی دارند. این به این معناست که آنها می توانند اطلاعات را در طول دنباله بهتر حفظ کنند. اما RNN ها معمولاً در درک وابستگی های طولانی دچار مشکل "از بین رفتن گرادیان" می شوند.

• Transformer ها از مکانیزم توجه استفاده می‌کنند که به مدل این امکان را می‌دهد که به طور دقیق تر و از اطلاعات موجود در دنباله ورودی استفاده کند.

معایب:

- مدل‌های Transformer معمولاً نیاز به حافظه بالاتری نسبت به RNN ها دارند، به ویژه زمانی که با دنباله های بلند و یا با تعداد پارامترهای زیادی مواجه هستند.
- Transformer ها معمولاً پیچیدگی محاسباتی بالایی دارند، که این می‌تواند زمان آموزش و استفاده از این مدل ها را افزایش دهد.

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

کدگذاری موقعیتی (positional encoding) به هر توکن در دنباله ورودی، یک بردار عددی اختصاص می‌دهد که بیانگر موقعیت آن توکن در دنباله است. (برای مثال، کلمه اول جمله ممکن است بردار $[1, 0, 0]$ را داشته باشد، کلمه دوم بردار $[0, 1, 0]$ و کلمه سوم بردار $[0, 0, 1]$ را داشته باشد.) این بردارها به ورودی های اولیه مدل Transformer اضافه می‌شوند تا مدل بتواند اطلاعات ترتیبی را در طول پردازش درک کند.

درک ترتیب کلمات در یک جمله برای فهم معنی آن ضروری است. به عنوان مثال، جمله "گربه موش را گرفت" معنای متفاوتی با جمله "موش گربه را گرفت" دارد. مدل های مبتنی بر Transformer بدون Positional Encoding قادر به تمایز بین این دو جمله نخواهند بود زیرا برعکس مدل های RNN به صورت ذاتی اطلاعات ترتیبی را حفظ نمی‌کنند.

شرایط: positional encoding ها باید قابل تکرار و نا محدود باشند و همچنین طول مشخصی داشته باشند. برای این کار از توابع سینوسی و کسینوسی با فرکانس های مختلف استفاده میشود. همچنین در حالت ایده آل، معیارهای زیرباید برآورده شوند:

فاصله بین هر دو گام زمانی در جمله هایی با طول های مختلف، یکسان باشد- برای هر گام زمانی (موقعیت کلمه در یک جمله) یک کد منحصر به فرد تولید کند- بتواند به جمله های طولانی تر تعمیم بدهد- قطعی باشد و مقادیر محدود داشته باشد.

(۶)

1. Encoder-only:

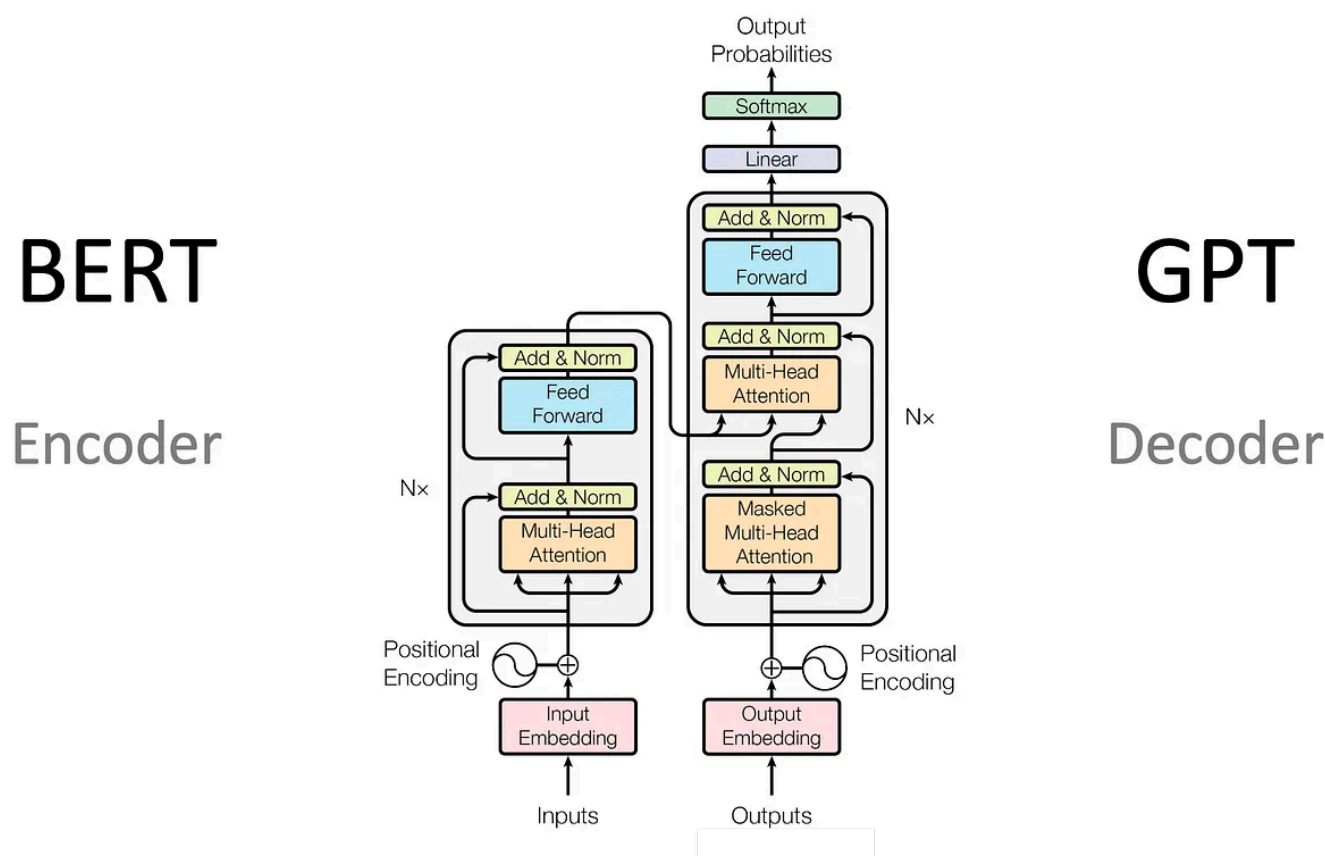
در این معماری، مدل فقط از بخش کدگذار تشکیل شده است. کدگذار دنباله ورودی را (برای مثال، یک جمله) پردازش می‌کند و نمایندگی های متناظر برای هر توکن در دنباله تولید می‌کند. این معماری برای وظایفی کاربرد دارد که تنها درک دنباله ورودی ضروری است، مانند طبقه بندی متن یا استخراج ویژگی ها.

2. Decoder-only:

در یک معماری کدگشا فقط، مدل فقط از بخش کدگشا تشکیل شده است. کدگشا دنباله ای از توکن ها (معمولاً توکن های شروع) را به عنوان ورودی دریافت کرده و به صورت خودکار یک دنباله خروجی را تولید می‌کند. این معماری معمولاً برای وظایفی مانند تولید متن، ترجمه ماشینی و مدل سازی زبان، که نیاز به تولید دنباله بر اساس نمایندگی های یادگرفته شده دارند، استفاده می‌شود.

:Encoder-Decoder.3

معماری کدگذار-کدگشا هر دو بخش کدگذار و کدگشا را ترکیب می‌کند. کدگذار دنباله ورودی را پردازش می‌کند و نمایندگی‌های متناظر را تولید می‌کند، در حالی که کدگشا بر اساس این نمایندگی‌ها یک دنباله خروجی تولید می‌کند. این معماری برای وظایفی مانند ترجمه ماشینی، خلاصه‌سازی متن و پاسخ به سوالات استفاده می‌شود، که دنباله‌های ورودی و خروجی ممکن است اندازه و ساختار متفاوتی داشته باشند.



(۷)

سیستم‌های استخراجی پاسخی را از یک مجموعه متن مرجع استخراج می‌کنند و برای انجام این کار، از تکنیک‌هایی مانند تطبیق الگو و استخراج کلیدواژه استفاده می‌کنند. اما سیستم‌های انتزاعی پاسخی را از ابتدا تولید می‌کنند که ممکن است در متن مرجع نباشد. این سیستم‌ها به جای جستجوی متن برای پاسخ‌های موجود، از درک خود از سوال و دانش دنیای واقعی برای تولید پاسخی جدید استفاده می‌کنند. برای انجام این کار، از تکنیک‌هایی مانند پردازش زبان طبیعی و استدلال استفاده می‌کنند.

همچنین به طور کلی می‌توان گفت برای پاسخ دادن به سوالات حالت‌های مختلفی داریم از جمله abstractive یا انتزاعی، extractive یا استخراجی، چند گزینه‌ای و یا بله/خیر. (توضیحات بیشتر در عکس زیر برگرفته از <https://arxiv.org/pdf/2005.00700> است)

Extractive [SQuAD]

Question: At what speed did the turbine operate?

Context: (Nikola_Tesla) On his 50th birthday in 1906, Tesla demonstrated his 200 horsepower (150 kilowatts) 16,000 rpm bladeless turbine. ...

Gold answer: 16,000 rpm

Abstractive [NarrativeQA]

Question: What does a drink from narcissus's spring cause the drinker to do?

Context: Mercury has awakened Echo, who weeps for Narcissus, and states that a drink from Narcissus's spring causes the drinkers to "Grow dotingly enamored of themselves." ...

Gold answer: fall in love with themselves

Multiple-Choice [ARC-challenge]

Question: What does photosynthesis produce that helps plants grow?

Candidate Answers: (A) water (B) oxygen (C) protein (D) sugar

Gold answer: sugar

Yes/No [BoolQ]

Question: Was America the first country to have a president?

Context: (President) The first usage of the word president to denote the highest official in a government was during the Commonwealth of England ...

Gold answer: no

(۸

می توان گفت که با Multiple head attention ها مدل وابستگی بین کلمات را بهتر می آموزد. برای مثال مدل زیر تشخیص داده است که فعل was به I باز می گردد.

Select model

bert-base-cased

Input Sentence

there was a time and I was alone, nowhere to go and no place to call home

Filters

Hide Special Tokens ☒

Show top 70% of att:

Layer

1 2 3 4 5 6 7 8 9 10 11 12

Selected heads:

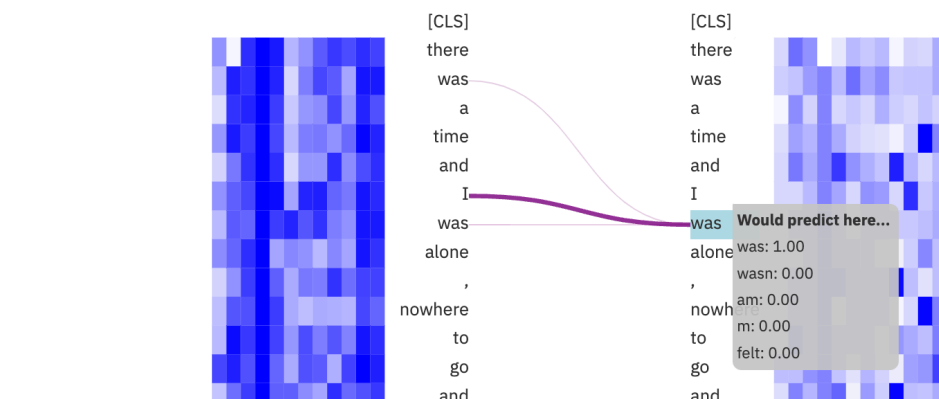
1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12

Select all heads

Unselect all heads

You focus on one token by **click**. You can mask any token by **double click**.

You can select and de-select a head by a **click** on the heatmap columns



Resources :

<https://www.youtube.com/watch?v=ZMxVe-HK174>

<https://machinelearningmastery.com/a-gentle-introduction-to-positional-encoding-in-transformer-models-part-1/>

<https://www.quora.com/Natural-Language-Processing-What-is-Non-factoid-question-answering#:~:text=A%20non%2Dfactoid%20question%20answering,and%20so%20on%2C%20so%20forth.>

[https://pure.port.ac.uk/ws/portalfiles/portal/59310424/Factoid vs. Non factoid Question Identification.pdf](https://pure.port.ac.uk/ws/portalfiles/portal/59310424/Factoid_vs._Non_factoid_Question_Identification.pdf)

<https://www.innerdoc.com/periodic-table-of-nlp-tasks/70-question-answering/#:~:text=Extractive%20QA%20has%20the%20goal,substring%20of%20the%20reference%20text>

<https://whites.agency/blog/open-domain-question-answering-introduction-to-the-topic/>

<https://medium.com/@minh.hoque/a-comprehensive-overview-of-transformer-based-models-encoders-decoders-and-more-e9bc0644a4e5>