



دانشگاه اصفهان
دانشکده مهندسی کامپیوتر

توسعه یک مدل زبانی پزشکی مبتنی بر استدلال در زبان فارسی

پایان نامه برای دریافت درجه کارشناسی ارشد در رشته هوش مصنوعی و رباتیکز

مهرداد قصایی

استاد راهنما

دکتر حمیدرضا برادران کاشانی

۱۴۰۴ آذر

بِسْمِ اللّٰهِ الرَّحْمٰنِ الرَّحِيْمِ



دانشگاه اصفهان
دانشکده مهندسی کامپیوتر

توسعه یک مدل زبانی پزشکی مبتنی بر استدلال در زبان فارسی

پایان نامه برای دریافت درجه کارشناسی ارشد در رشته هوش مصنوعی و رباتیکز

مهرداد قصابی

استاد راهنما

دکتر حمیدرضا برادران کاشانی

۱۴۰۴ آذر



گواهی دفاع از پایان نامه کارشناسی ارشد

هیات داوران پایان نامه کارشناسی ارشد آقای مهرداد قصابی به شماره دانشجویی ۴۰۲۳۶۱۴۰۲۹ در رشته هوش مصنوعی را در تاریخ با عنوان « توسعه یک مدل زبانی پزشکی مبتنی بر استدلال در زبان فارسی »
به عدد **_____** به حروف **_____**

_____	_____
-------	-------

با نمره نهایی

ارزیابی کرد.

و درجه

ردیف	مشخصات هیات داوران	نام و نام خانوادگی	مرتبه دانشگاهی	دانشگاه یا موسسه	امضا
۱	استاد راهنما	دکتر حمیدرضا برادران کاشانی	استاد	دانشگاه اصفهان	
۲	استاد داور داخلی	دکتر داور داخلی	دانشیار	دانشگاه اصفهان	
۳	استاد مدعو	دکتر داور خارجی	دانشیار	دانشگاه اصفهان	
۴	نماینده تحصیلات تکمیلی دانشکده	دکتر نماینده	دانشیار	دانشگاه اصفهان	

نام و نام خانوادگی معاون آموزشی و تحصیلات
تکمیلی پرdisیس دانشکده های فنی:

نام و نام خانوادگی معاون تحصیلات تکمیلی و
پژوهشی دانشکده / گروه:

تاریخ و امضا:

تاریخ و امضا:

تعهدنامه اصالت اثر

به نام خدا

اینجانب مهرداد قصابی تأیید می‌کنم که مطالب مندرج در این پایان‌نامه حاصل کار پژوهشی اینجانب است و به دستاوردهای پژوهشی دیگران که در این نوشه از آن‌ها استفاده شده است مطابق مقررات ارجاع گردیده است. این پایان‌نامه قبل^ا برای احراز هیچ مدرک هم‌سطح یا بالاتری ارائه نشده است.

نام و نام خانوادگی دانشجو: مهرداد قصابی

تاریخ و امضای دانشجو:

کلیه حقوق مادی و معنوی این اثر
متعلق به دانشگاه اصفهان است.

تقدیم به:

پدرم که در طول تحصیل پشتیبانم بوده است

قدردانی

دستان پدر و مادر نازنینم را به پاس مهر بیکرانشان به گرمی میفسارم و از استاد راهنمای خود جناب آفای دکتر حمیدرضا برادران بابت راهنمایی هایشان در طول انجام این پایان نامه سپاس گزاری میکنم.

مهرداد قصابی

۱۴۰۴ آذر

چکیده

استفاده از هوش مصنوعی در پاسخگویی به سوالات پزشکی به عنوان یکی از حوزه‌های نوظهور و مهم در فناوری و بهداشت شناخته می‌شود که در سال‌های اخیر مورد توجه گستردگانی قرار گرفته است. این فناوری پیشرفته، با قابلیت‌های ویژه خود، می‌تواند کیفیت خدمات پزشکی ارائه شده به بیماران را به شکل چشمگیری ارتقا دهد. همچنین، با سرعت بخشیدن به فرآیند ارائه اطلاعات پزشکی و ارائه پاسخ‌های سریع و دقیق به سوالات پزشکان و بیماران، نقش مهمی در کاهش فشار کاری پزشکان ایفا می‌کند. به این ترتیب، هوش مصنوعی نه تنها موجب افزایش کارایی در سیستم‌های بهداشتی می‌شود، بلکه تجربه کلی بیماران را بهبود می‌بخشد و زمینه ارائه درمان‌های بهتر و مؤثرتر را فراهم می‌کند.

از طرف دیگر از آنجا که پزشکی مبتنی بر استدلال و تحلیل‌های منطقی است، توسعه یک مدل پزشکی که بر پایه زنجیره‌ای از افکار و استدلال‌های منطقی طراحی شده باشد، می‌تواند دقت و کارایی این مدل را به طور قابل توجهی افزایش دهد. چنین رویکردی امکان انجام فرآیندهای پیچیده تشخیصی و درمانی را به صورت ساختاریافته‌تر و هدفمندتر فراهم می‌کند. در این زمینه، هر مرحله از تشخیص و درمان باید مبتنی بر شواهد علمی و داده‌های معتبر باشد. به عنوان مثال، پزشکان در فرآیند تشخیص بیماری‌ها معمولاً از تاریخچه پزشکی، علائم بالینی و نتایج آزمایش‌ها بهره می‌گیرند. با طراحی یک مدل منطقی، این داده‌ها می‌توانند در قالب یک زنجیره منطقی به یکدیگر متصل شوند که به شناسایی الگوهای روابط میان علائم و بیماری‌ها کمک می‌کند.

واژگان کلیدی هوش مصنوعی در پزشکی، مدل‌های زبانی فارسی، مدل‌های زبانی پزشکی، پردازش زبان‌های طبیعی، توانایی استدلال هوش مصنوعی

فهرست مطالب

ج	فهرست تصاویر
چ	فهرست جداول
ح	فهرست الگوریتم‌ها
د	فهرست برنامه‌ها
۱	فصل ۱: دیباچه
۱	۱.۱ هدف پژوهش
۱	۲.۱ کاربرد پژوهش
۱	۱.۲.۱ کاربرد مدل‌های زبانی پزشکی
۲	۲.۲.۱ کاربرد مدل‌های زبانی پزشکی دارای قابلیت استدلال
۲	۳.۲.۱ کاربرد مدل‌های زبانی پزشکی فارسی
۳	۳.۱ مراحل انجام پایان نامه
۴	۴.۱ ساختار پایان نامه
۵	فصل ۲: ادبیات موضوع
۵	۱.۲ مقدمه
۵	۲.۲ مدل‌های زبانی
۶	۱.۲.۲ مدل‌های زبانی آماری
۶	۲.۲.۲ مدل‌های زبانی بازگشتنی

۷	۳.۲.۲	مدل های زبانی مبتنی بر ترنسفورمر
۷	۱.۳.۲.۲	مدل های زبانی فقط رمزگذار
۸	۲.۳.۲.۲	مدل های زبانی فقط رمزگشا
۸	۳.۳.۲.۲	مدل های زبانی رمزگذار-رمزگشا
۸	۳.۲	سیستم های پرسش و پاسخ
۹	۱.۳.۲	سیستم های پرسش و پاسخ استخراجی
۹	۲.۳.۲	سیستم های پرسش و پاسخ تولیدی
۹	۴.۲	شیوه های سنجش مدل های زبانی
۱۰	۱.۴.۲	سنجش بر اساس میزان پاسخگویی به پرسش و پاسخ های چند گزینه ای
۱۰	۲.۴.۲	سنجش بر اساس نظر مدل داور
۱۱	۳.۴.۲	سنجش بر اساس استنتاج زبان طبیعی
۱۱	۴.۴.۲	سنجش بر اساس امتیاز bert
۱۲	۵.۲	یادگیری تقویتی در مدل های زبانی
۱۲	۱.۵.۲	یادگیری تقویتی با بازخورد انسان یا هوش مصنوعی
۱۳	۲.۵.۲	الگوریتم های بهینه سازی مدل زبانی

۱۷	فصل ۳: بررسی کارهای پیشین
۱۷	۱.۳ مقدمه
۱۷	۲.۳ کارهای پیشین در حوزه زبان انگلیسی
۱۷	۱.۲.۳ مدل های MedPaLM
۱۸	۲.۲.۳ مدل ChatDoctor
۱۸	۳.۲.۳ مدل های Meerkat
۱۹	۴.۲.۳ مدل MedMobile
۱۹	۳.۳ کارهای پیشین در حوزه زبان فارسی
۲۰	۱.۳.۳ مدل Sina-bert
۲۰	۲.۳.۳ سیستم پرسش و پاسخ پزشکی دکتر ویسی و همکاران

۲۰	پایان نامه کارشناسی ارشد خانم لیلا دارابی	۳.۳.۳
۲۱	بهبود قابلیت دلیل آوری در حوزه پزشکی	۴.۳
۲۱ MedSS	۱.۴.۳
۲۲ MedReason	۲.۴.۳
۲۳ HuatuoGPT-o1	۳.۴.۳
۲۳	بهبود قابلیت دلیل آوری در حوزه های دیگر	۵.۳
۲۳ Deepseek	۱.۵.۳
۲۴ Thought Preference Optimization	۲.۵.۳
۲۵ چهارچوب آموزش استدلال مدل های بزرگ به مدل های کوچک	۳.۵.۳
فصل ۴: جمع آوری دادگان		
۲۷ مقدمه	۱.۴
۲۷ معرفی پیکره پزشکی فارسی	۲.۴
۲۸ MF3QA	۳.۴
۲۹ MF3QA	۱.۳.۴
۲۹ MF3QA	۲.۳.۴
۳۱ خزش از تالار گفتگو دکترهست	۱.۲.۳.۴
۳۱ ترجمه قسمت پزشکی مجموعه داده MMLU	۴.۴
۳۲ گردآوری سوالات کنکور علوم پایه پزشکی ایران	۵.۴
۳۳ MedMCQA	۶.۴
۳۳ PersianMedQA و PerMedCQA	۷.۴
فصل ۵: معرفی مدل زبانی گائوکرنا-V		
۳۵ مقدمه	۱.۵
۳۵ مدل پایه	۲.۵
۳۶ aya-expanses	۱.۲.۵
۳۷ تنظیم دقیق روی پیکره پزشکی	۳.۵

۴.۵	تنظیم دستورالعملی روی مجموعه داده MF3QA	۳۸
۵.۵	رد پای کربن مدل گائوکرنا-V	۳۸
۶.۵	نتایج	۲۹
۱.۶.۵	مقایسه با مدل های زبانی فارسی همه منظوره	۴۰
۲.۶.۵	مقایسه با جایگزین های خط لوله ای	۴۱
فصل ۶: استدلال در مدل های زبانی		۴۵
۱.۶	ناتوانی معماری ترانسفورمر در تفکر سیستم دو	۴۵
۲.۶	طراحی یک معماری دارای قابلیت تفکر سیستم دو	۴۶
۳.۶	بهبود توانایی دلیل آوری مدل های زبانی مبتنی بر ترانسفورمر	۴۷
۱.۳.۶	روش های مبتنی بر هوش مصنوعی عصبی-نمادین	۴۷
۲.۳.۶	روش های مبتنی بر راهنمایی مدل زبانی توسط یک مدل پاداش	۴۹
۳.۳.۶	روش های مبتنی بر تمرین روی داده های دلیل آوری	۴۹
فصل ۷: معرفی مدل زبانی گائوکرنا-R		۵۱
۱.۷	مقدمه	۵۱
۲.۷	مدل پایه	۵۱
۳.۷	متدهای معرفی شده	۵۲
۱.۳.۷	روش نخست	۵۲
۲.۳.۷	روش دوم	۵۳
۴.۷	رد پای کربن	۵۴
۵.۷	نتایج	۵۵
۱.۵.۷	سنجهش استدلال پزشکی	۵۶
۱.۱.۵.۷	دقت	۵۷
۲.۱.۵.۷	Pass@k	۵۷
۲.۵.۷	سنجهش دانش پزشکی	۵۸
۳.۵.۷	سنجهش پایانی	۵۹

۶۵

فصل ۸: نتیجه گیری

۶۷

کتاب نامه

فهرست تصاویر

۱.۲	معماری ترانسفورمر	۱۵
۲.۲	مقایسه الگوریتم بهینه‌سازی سیاست مجاورتی و بهینه‌سازی مستقیم ترجیحات	۱۵
۱.۳	بلوک دیاگرام مدل medsss	۲۲
۲.۳	بلوک دیاگرام Thought Preference Optimization	۲۵
۳.۳	بلوک دیاگرام آموزش دلیل آوری توسط مدل بزرگتر به مدل کوچکتر	۲۶
۱.۴	سهم هر مجله در پیکره پزشکی فارسی گردآوری شده	۳۰
۲.۴	سهم هر تالار گفتگو در مجموعه داده MF3QA	۳۱
۱.۵	mekanisem آربیتاژ داده	۳۶
۲.۵	نمودار کاهش خط تنظیم دقیق روی پیکره پزشکی	۳۸
۳.۵	نمودار کاهش خط تنظیم دقیق روی مجموعه داده MF3QA	۳۹
۴.۵	mekanisem جایگزین خط لوله ای	۴۰
۵.۵	نرخ پیروزی گائوکرنا-V در رقابت با بقیه مدل های زبانی فارسی همه منظوره	۴۰
۶.۵	نرخ پیروزی گائوکرنا-V در رقابت با جایگزین های خط لوله ای	۴۲
۱.۷	بلوک دیاگرام شیوه نخست	۵۳
۲.۷	بلوک دیاگرام شیوه دوم	۵۵
۳.۷	نمودار کاهش خط بهینه سازی مستقیم ترجیحات	۵۷
۴.۷	نتایج Pass@k روی مجموعه داده ترجمه قسمت پزشکی MMLU	۵۹
۵.۷	نتایج Pass@k روی مجموعه داده کنکور علوم پایه پزشکی شهریور ۱۴۰۱	۶۰

فهرست جداول

۳	اطلاعات دوفاز پایان نامه	۱.۱
۲۸	مقایسه پیکره گردآوری شده با پیکره های گردآوری شده توسط I. Garcia Ferrero et al.	۱.۴
۲۹	مقایسه مجموعه داده های پرسش و پاسخ آزاد پزشکی با مجموعه داده گردآوری شده . . .	۲.۴
۴۳	مقایسه مدل گائوکرنا-۷ با بقیه مدل های زبانی فارسی همه منظوره	۱.۵
۴۴	مقایسه مدل گائوکرنا-۷ با جایگزین های خط لوله ای	۲.۵
۶۱	دقت با درخواست زنجیره افکار بدون نمره منفی	۱.۷
۶۲	دقت با درخواست زنجیره افکار با نمره منفی	۲.۷
۶۳	دقت با درخواست مستقیم	۳.۷
۶۴	ارزیابی دوپیکربندی	۴.۷

فهرست الگوریتم‌ها

۳۲	الگوریتم جستجو اول عرض برای استخراج رکورد های پرسش و پاسخ پزشکی	۱.۴
۵۴	شیوه دوم پیشنهاد شده برای بهبود دلیل آوری	۱.۷
۵۶	شیوه دوم پیشنهاد شده برای بهبود دلیل آوری	۲.۷

فهرست برنامه‌ها

فصل ۱

دیباچه

۱.۱ هدف پژوهش

هدف از این پژوهش، توسعه یک مدل زبانی پزشکی فارسی بر پایه استدلال^۱ است که قابلیت اجرا روی دستگاه‌های محلی را داشته باشد. اجرا روی دستگاه‌های محلی از آن جهت حائز اهمیت است که داده‌های پزشکی اغلب حساس و خصوصی هستند و ارسال آنها به سرورهای خارجی ممکن است خطرات جدی برای حریم خصوصی بیماران ایجاد کند. [۱]

۲.۱ کاربرد پژوهش

۱.۲.۱ کاربرد مدل‌های زبانی پزشکی

مدل‌های زبانی در سال‌های اخیر با استفاده از داده‌های بسیار گسترده‌تر و معماری‌های پیشرفته‌تر به پیشرفت‌های چشمگیری دست یافته‌اند. این مدل‌ها توانایی درک بهتر مفاهیم، تولید متن‌های طبیعی‌تر و پاسخ‌دهی دقیق‌تر به سوالات را پیدا کرده‌اند.

این پیشرفت‌ها منجر به افزایش چشمگیر کاربرد هوش مصنوعی^۲ در حوزه‌های مختلف، بهویژه در زمینه

^۱ reasoning
^۲ artificial intelligence

پزشکی، شده است. امروزه در حوزه پزشکی، مدل‌های زبانی مبتنی بر یادگیری ژرف^۳ نقش مهمی در تحلیل داده‌های پزشکی، بهبود دقت تشخیص بیماری‌ها، ارائه پیشنهادهای درمانی دقیق‌تر و افزایش کیفیت مراقبت از بیماران ایفا می‌کنند. علاوه بر این، این فناوری، به بهینه‌سازی سیستم‌های اداری و کاهش بار کاری کادر درمانی کمک شایانی کرده است. به عنوان مثال، مدل‌های هوش مصنوعی قادرند با تحلیل داده‌های حاصل از پرونده‌های پزشکی، الگوهای مرتبط با بیماری‌ها را شناسایی کنند و اطلاعات ارزشمندی را برای تصمیم‌گیری سریع‌تر و دقیق‌تر در اختیار پزشکان قرار دهند. این مدل‌ها همچنین می‌توانند نقش مهمی در تکمیل مشاوره‌های پزشکی ایفا کرده و به پزشکان در ارائه اطلاعات دقیق‌تر و سریع‌تر کمک کنند. و حتی شاید در آینده ای نه چندان دور بتوانند جای پزشکان را در مشاوره‌های پزشکی بگیرند.

این تحول نه تنها به افزایش کارایی و بهره‌وری در سیستم‌های درمانی منجر شده است، بلکه تجربه کلی بیماران را نیز بهبود بخشدیده و امکان ارائه خدمات درمانی بهتر و مؤثرتر را فراهم کرده است. به همین دلیل، توسعه و استفاده از مدل‌های زبانی پزشکی^۴، همچنان مورد توجه پژوهشگران و متخصصان قرار دارد.

۲.۰.۱ کاربرد مدل‌های زبانی پزشکی دارای قابلیت استدلال

پزشکی به عنوان یک علم تجربی و کاربردی، بیش از هر چیز به استدلال دقیق وابسته است. تشخیص بیماری، انتخاب درمان، پیش‌بینی پیامد و حتی آموزش پزشکی، همگی بر پایه زنجیره‌های منطقی، تحلیل داده‌های بالینی و استنتاج از شواهد بنا شده‌اند. از آن روی که پزشکی علم وابسته به استدلال است، توسعه یک مدل زبانی که دارای قابلیت‌های استدلال برتری باشد می‌تواند بسیار تحول‌آفرین باشد. این مدل‌ها نه تنها می‌توانند اطلاعات را بازیابی کنند، بلکه می‌توانند دلیل بیاورند، احتمالات را مقایسه کنند، تناقضات را شناسایی کنند و تصمیمات را توجیه‌پذیر سازند. بنابراین توسعه این نوع از مدل‌های زبانی در حوزه پزشکی بسیار مفید است.

۳.۰.۱ کاربرد مدل‌های زبانی پزشکی فارسی

علیرغم پیشرفت‌های چشمگیر در توسعه مدل‌های زبانی پزشکی به زبان انگلیسی، در حوزه زبان فارسی هنوز کار چندانی صورت نگرفته است. این در حالی است که در سرتاسر جهان میلیون‌ها نفر تنها قادر به استفاده از این زبان هستند؛ بنابراین تلاش برای توسعه یک مدل زبانی پزشکی در زبان فارسی می‌تواند گامی رو به جلو در ارتباطات و خدمات درمانی کشورهای فارسی زبان باشد.

^۳ deep learning
^۴ medical language models

۳.۱ مراحل انجام پایان نامه

همانطور که در جدول ۱.۱ این پایان نامه در دو فاز اصلی طراحی و اجرا شده است، فاز نخست به جمع‌آوری دادگان پزشکی فارسی و توسعه مدلی با نام گائوکرنا-V اختصاص دارد که قادر توانایی استدلال بوده و بیشتر بر درک سیستم یک^۵ زبان تمرکز دارد. از این فاز، مقاله‌ای با عنوان "اهرم قرار دادن داده‌های آنلاین برای بهبود دانش پزشکی یک مدل زبانی کوچک پزشکی فارسی"^۶ استخراج شده است که به تشریح فرآیند جمع‌آوری داده‌ها و نحوه بهینه‌سازی دانش پزشکی مدل می‌پردازد. در فاز دوم این پژوهش ابتدا تکنیک‌های جدیدی برای ارتقای توانایی استدلال و درک سیستم دو مدل معرفی شده و سپس مدل گائوکرنا-R در این فاز توسعه داده شده است. از این فاز نیز، مقاله‌ای با عنوان "بهبود مهارت‌های استدلال در مدل‌های زبانی پزشکی فارسی کوچک می‌تواند از آموزش داده‌های بزرگ مقیاس پیشی بگیرد"^۷ استخراج شده است.

مقاله نخست درسی و دومین کنفرانس ملی و دهمین کنفرانس بین‌المللی مهندسی پزشکی ایران و مقاله دوم در یازدهمین کنفرانس بین‌المللی پردازش سیگنال و سیستم‌های هوشمند منتشر شده است.

	gaokerena-V	gaokerena-R
مخزن گیت هاب	mehrdadghassabi/gaokerena-V	mehrdadghassabi/gaokerena-R
مخزن پارامترها	gaokerena/gaokerena-v1.0	gaokerena/gaokerena-r1.0
پیوند مقاله	arxiv.org/pdf/2505.16000	arxiv.org/pdf/2510.20059
هزینه	۳۰۰ دلار	۷۰ دلار
توانایی استدلال	خیر	بله
منتشر شده در	سی و دومین کنفرانس ملی و دهمین کنفرانس بین‌المللی مهندسی پزشکی ایران	یازدهمین کنفرانس بین‌المللی پردازش سیگنال و سیستم‌های هوشمند

جدول ۱.۱: اطلاعات دو فاز پایان نامه

^۵ در علم رفتارشناسی به درک سریع، شهودی و بدون نیاز به تفکر ژرف درک سیستم یک و به درک آهسته، غیر شهودی و نیازمند استدلال درک سیستم دو می‌گویند.

Leveraging Online Data to Enhance Medical Knowledge in a Small Persian Language Model^۶
Enhancing Reasoning Skills in Small Persian Medical Language Models Can Outperform Large-Scale^۷
Data Training

۴.۱ ساختار پایان نامه

در این پایان نامه، ساختار فصل‌ها به گونه‌ای طراحی شده است که مراحل مختلف پژوهش به صورت منظم و هدفمند ارائه شوند. فصل دوم به بررسی کارهای پیشین اختصاص دارد که در آن مطالعات انجام‌شده در زمینه‌های مرتبط مرور خواهد شد. در فصل سوم، به دلیل عدم وجود دادگان پژوهشی در حوزه زبان فارسی، فرآیند گردآوری و آماده‌سازی این دادگان به‌طور دقیق تشریح خواهد شد. سپس در فصل چهارم، با استفاده از دادگان معرفی‌شده در فصل سوم، مدل اولیه با نام گائوکرنا-V^۸ معرفی و تحلیل می‌شود. در فصل پنجم، توانایی‌های استدلال در مدل‌های هوش مصنوعی مورد بررسی قرار گرفته و چالش‌ها و راهکارهای مرتبط با استدلال ارائه خواهند شد. در ادامه، در فصل ششم، با معرفی تکنیک‌هایی برای بهبود توانایی استدلال یک مدل زبانی مدل پیشرفته‌تری به نام گائوکرنا-R معرفی و ویژگی‌های آن به تفصیل شرح داده می‌شود. در نهایت، فصل پایانی به جمع‌بندی نتایج پژوهش و پیشنهاداتی برای تحقیقات آینده اختصاص دارد.

^۸ نام گائوکرنا از درختی افسانه‌ای الهام گرفته شده است که در روایات اساطیری زرتشتی به عنوان نماد شفاده‌ی و جاودانگی شناخته می‌شود.

فصل ۲

ادبیات موضوع

۱.۲ مقدمه

در این فصل به بررسی مفاهیم مدل‌های زبانی^۱ انواع آن و چگونگی سنجش آن‌ها خواهیم پرداخت. همچنین، انواع سیستم‌های پرسش و پاسخ^۲ را مورد تحلیل قرار خواهیم داد تا درک بهتری از عملکرد و کاربردهای مختلف این سیستم‌ها به دست آوریم.

۲.۲ مدل‌های زبانی

مدل‌های زبانی ابزارهای پیشرفته‌ای هستند که برای پردازش و تولید زبان طبیعی طراحی شده‌اند. این مدل‌ها با استفاده از یادگیری ماشین^۳ و به ویژه یادگیری عمیق^۴، توانایی درک و تولید متن را دارند. در حقیقت، وظیفه مدل‌های زبانی پیش‌بینی توکن بعدی^۵ بر اساس متنی است که تاکنون تولید شده است. این مدل‌ها دارای انواع مختلفی است که در ادامه آنها را بررسی می‌کنیم.

language models^۱
question answering systems^۲
machine learning^۳
deep learning^۴
next token prediction^۵

۱.۲.۲ مدل‌های زبانی آماری

مدل‌های زبانی n-gram ابتدایی ترین مدل‌های زبانی هستند که در دهه نود میلادی به عنوان جایگزینی برای ترجمه ماشینی مبتنی بر قانون^۶ معرفی شدند [۲] همانطور که در فرمول ۱.۲.۲ مشاهده می‌کنید این مدل‌ها برای پیش‌بینی توکن بعدی از فراوانی n-gram ها در پیکره^۷ موجود استفاده می‌کنند. به عنوان مثال در مدل زبانی ۲-gram، احتمال وقوع یک کلمه تنها بر اساس کلمه قبلی محاسبه می‌شود.

$$P(w_1, w_2, \dots, w_n) = \prod_{i=1}^n P(w_i | w_{i-n+1}, w_{i-n+2}, \dots, w_{i-1})$$

۲.۲.۲ مدل‌های زبانی بازگشتی

شبکه‌های عصبی بازگشتی نوعی از شبکه‌های عصبی مصنوعی^۸ هستند که به طور خاص برای پردازش توالی‌ها^۹ طراحی شده‌اند. [۳] از آنجاکه زبان را نیز می‌توان به صورت یک توالی از توکن‌ها تعریف کرد، بنابراین می‌توان از شبکه‌های عصبی بازگشتی به عنوان مدل‌های زبانی استفاده کرد.

همانطور که در فرمول ۲.۲.۲ که فرمول پایه شبکه‌های عصبی است می‌بینید ویرگی اصلی شبکه‌های عصبی بازگشتی این است که دارای حلقه‌های بازگشتی است که به آن اجازه می‌دهد اطلاعات را از مراحل قبلی به مراحل بعدی منتقل کند. این ساختار باعث می‌شود شبکه‌های عصبی بازگشتی بتواند وابستگی‌های زمانی و ترتیبی داده‌ها را مدل‌سازی کند.

شبکه‌های عصبی بازگشتی مانند نوع ساده آن^{۱۰} معمولاً در مدل کردن وابستگی‌های بلند مدت^{۱۱} با مشکل مواجه می‌شوند، این وابستگی‌ها که در زبان‌های طبیعی به وفور یافت می‌شوند باعث شده‌اند که مدل‌های زبانی مبتنی بر شبکه‌های عصبی بازگشتی مدل‌های چندان خوبی نباشند. هر چند در انواع دیگر این شبکه‌ها مانند LSTM [۴] برای حل این مشکل تلاش شده است اما این مشکل هنوز در این نوع از شبکه‌های عصبی وجود دارد.

$$h_t = f(W_h h_{t-1} + W_x x_t + b)$$

rule-based machine translation ^۹	
corpus ^۷	
artificial neural networks ^۸	
sequences ^۹	
vanilla recurrent neural networks ^{۱۰}	
long term dependencies ^{۱۱}	

۳.۰.۲.۲ مدل‌های زبانی مبتنی بر ترنسفورمر

همان طور که در قسمت پیشین اشاره شد، مدل‌های زبانی بازگشتی در مدل‌سازی وابستگی‌های بلندمدت با مشکل مواجه هستند و به دلیل ماهیت توالی گونه خود، سرعت پردازش آنها نیز بسیار پایین است.

مقاله Attention is All You Need [۵] با معماری نوآورانه خود توانست هر دوی این مشکلات را حل کند. این معماری از مکانیزم توجه^{۱۲} استفاده می‌کند که به مدل اجازه می‌دهد تا به طور همزمان به تمامی ورودی‌ها توجه کند و وابستگی‌های بلندمدت را به راحتی شناسایی کند. به این ترتیب، سرعت پردازش به طور قابل توجهی افزایش می‌یابد.

علاوه بر این، معماری ترنسفورمر^{۱۳} که در شکل ۱.۲ میتوانید آن را ببینید، قابلیت پردازش موازی را دارد، زیرا معماری آن ماهیت توالی گونه ندارد. به همین دلیل، با استفاده از پردازنده‌های گرافیکی^{۱۴} می‌توان پردازش‌های موازی را انجام داد که سرعت پردازش را به طور چشمگیری بهبود می‌بخشد.

مدل‌های مبتنی بر معماری ترنسفورمر سه دسته هستند که در ادامه به بررسی این سه دسته خواهیم پرداخت.

۱.۳.۰.۲.۲ مدل‌های زبانی فقط رمزگذار

مدل‌های زبانی فقط رمزگذار^{۱۵}، تنها از بخش رمزگذار معماری ترنسفورمر که در سمت چپ شکل ۱.۲ قابل مشاهده است، استفاده می‌کنند. این مدل‌ها به طور ویژه برای پردازش و درک متن طراحی شده‌اند و ورودی‌ها را به یک نمایش داخلی تبدیل می‌کنند که شامل اطلاعات معنایی و ساختاری متن است.

این نوع مدل‌ها عمدتاً در وظایفی مانند تحلیل احساسات^{۱۶}، دسته‌بندی متن^{۱۷} و استخراج ویژگی‌ها^{۱۸} کاربرد دارند. آن‌ها به خوبی می‌توانند الگوهای زبانی و معنایی را شناسایی کرده و اطلاعات مفیدی از داده‌های متنی استخراج کنند، اما به تنهایی توانی تولید متن جدید را ندارند.

به عنوان مثال، می‌توان از BERT [۶]^[۷]، که یکی از معروف‌ترین مدل‌های زبانی فقط رمزگذار است، یاد کرد.

Attention Mechanism ^{۱۲}
transformer ^{۱۳}
graphical processing unit ^{۱۴}
encoder only language models ^{۱۵}
sentiment analysis ^{۱۶}
text classification ^{۱۷}
feature extraction ^{۱۸}

۲.۳.۲.۲ مدل‌های زبانی فقط رمزگشا

مدل‌های زبانی فقط رمزگشا^{۱۹} نوعی از مدل‌های زبانی هستند که به طور خاص برای تولید متن و پیش‌بینی توکن‌های بعدی در یک توالی طراحی شده‌اند. این مدل‌ها تنها از ساختار رمزگشای معماری ترانسفورمر که در سمت راست شکل ۱.۲ قابل مشاهده است، استفاده می‌کنند و به صورت تک‌جهته عمل می‌کنند، به این معنا که برای تولید هر توکن، تنها به توکن‌های قبلی خود در توالی دسترسی دارند.

در این مدل‌ها، هدف اصلی پیش‌بینی توکن بعدی بر اساس توکن‌های قبلی است. به عنوان مثال، اگر ورودی مدل یک جمله باشد، مدل سعی می‌کند کلمه بعدی را پیش‌بینی کند. این نوع از مدل‌ها در کاربردهایی مانند تولید متن، چت‌بات‌ها و ترجمه ماشینی بسیار موثر هستند.

مدل‌های رمزگشا معمولاً با استفاده از داده‌های متنه بزرگ آموزش دیده و توانایی بالایی در تولید متن‌های معنادار و مرتبط دارند. یکی از معروف‌ترین نمونه‌های این دسته از مدل‌ها، مدل GPT^{۲۰} [۸] است که توسط OpenAI توسعه یافته است.

۳.۳.۲.۲ مدل‌های زبانی رمزگذار-رمزگشا

دل‌های زبانی رمزگذار-رمزگشا^{۲۱} نوعی از مدل‌های زبانی هستند که هر دو قسمت معماری ترانسفورمر که در شکل ۱.۲ مشاهده می‌کنید استفاده می‌کند.

این معماری که در مدل هایی مانند T5 [۹] به کار گرفته شده در وظایف پیچیده‌ای مانند ترجمه ماشینی، خلاصه‌سازی متن و تولید گفتار کاربرد دارد. به عنوان مثال، در ترجمه ماشینی، بخش رمزگذار جمله‌ای را به زبان مبدأ تحلیل کرده و آن را به یک نمایش معنایی تبدیل می‌کند، سپس بخش رمزگشا این نمایش را به زبان مقصد ترجمه می‌کند.

۳.۲ سیستم‌های پرسش و پاسخ

سیستم‌های پرسش و پاسخ^{۲۲} فناوری‌های هوشمندی هستند که با استفاده از تکنیک‌های پردازش زبان طبیعی^{۲۳}، به کاربران این امکان را می‌دهند تا سوالات خود را مطرح کرده و پاسخ‌های دقیق و مرتبط دریافت کنند. این

Decoder-Only Language Models^{۱۹}
Generative Pre-trained Transformer^{۲۰}
Encoder-Decoder Language Models^{۲۱}
Question Answering Systems^{۲۲}
Natural Language Processing^{۲۳}

سیستم‌ها بر دو نوع هستند استخراجی^{۲۴} و تولیدی^{۲۵} که در ادامه به بررسی آنها خواهیم پرداخت. [۱۰]

۱.۳.۲ سیستم‌های پرسش و پاسخ استخراجی

سیستم‌های پرسش و پاسخ استخراجی به منظور پاسخ‌گویی به سوالات کاربران، به جستجوی اطلاعات در پایگاه‌های داده یا مستندات می‌پردازند و پاسخ‌ها را از متن استخراج می‌کنند. این سیستم‌ها معمولاً از یک مدل زبانی فقط رمزگذار مانند bert استفاده می‌کنند.

۲.۳.۲ سیستم‌های پرسش و پاسخ تولیدی

با توسعه و گسترش هوش مصنوعی تولید کننده^{۲۶} سیستم‌های پرسش و پاسخ‌هایی پدید آمدند که در آن برای پاسخ دادن به پرسش کاربر به دانش مدل زبانی تکیه می‌شود. یعنی مدل زبانی با توجه به دانشی که در زمینه پرسش مطرح شده دارد بایستی پاسخ را تولید کند. از آنجایی که در این سیستم‌ها بایستی چیزی تولید شود بنابراین در آن‌ها از مدل‌های زبانی فقط رمزگشا یا مدل‌های زبانی رمزگذار-رمزگشا استفاده می‌گردد.

با بهبود دانش پزشکی و توانایی استدلال یک مدل پایه مانیز در پایان نامه حاضر اقدام به طراحی یک سیستم پرسش و پاسخ تولیدی کرده ایم.

۴.۲ شیوه‌های سنجش مدل‌های زبانی

سنجش دانش یک مدل زبانی به ویژه در زمینه پزشکی از اهمیت بالایی برخوردار است، زیرا این فرآیند به شناخت ما از میزان دانش یک مدل زبانی، کمک شایانی می‌کند. در ادامه چندین روش سنجش کیفیت پاسخ‌های مدل‌های زبانی را مطرح خواهیم کرد، ضمناً ما از روش اول و دوم برای سنجش دانش مدل‌های خود استفاده کرده ایم.

^{۲۴} extractive

^{۲۵} generative

^{۲۶} intelligence artificial generative

۱.۴.۲ سنجش بر اساس میزان پاسخگویی به پرسش و پاسخ‌های چندگزینه‌ای

یکی از معیارهای مهم برای سنجش عملکرد مدل‌های زبانی، ارزیابی توانایی آن‌ها در پاسخگویی به پرسش‌های چهارگزینه‌ای است که از پیش آماده شده‌اند. این نوع ارزیابی به دلیل ساختار مشخص و استاندارد پرسش‌ها، امکان مقایسه دقیق‌تری بین مدل‌های مختلف را فراهم می‌آورد. یکی از مجموعه‌های داده‌ای که به طور گسترده در این زمینه مورد استفاده قرار می‌گیرد، مجموعه داده MMLU^{۲۷} [۱۱] است. این مجموعه شامل پرسش‌های متنوعی است که درباره موضوعات مختلفی مانند علوم، پژوهشی، ریاضیات، تاریخ، و ادبیات طراحی شده‌اند.

مجموعه داده MMLU به عنوان یک استاندارد در ارزیابی مدل‌های زبانی، به محققان و توسعه‌دهندگان این امکان را می‌دهد که عملکرد مدل‌های زبانی خود را در زمینه‌های مختلف بسنجند و نقاط قوت و ضعف آن‌ها را شناسایی کنند. پرسش‌های چهارگزینه‌ای در MMLU به گونه‌ای طراحی شده‌اند که نیاز به درک عمیق و تحلیل دقیق متن دارند. این ویژگی، مدل‌های را به چالش می‌کشد تا نه تنها اطلاعات را بازیابی کنند، بلکه توانایی استدلال و تحلیل خود را نیز به نمایش بگذارند. با استفاده از این معیار، می‌توان به راحتی مقایسه‌هایی بین مدل‌های مختلف انجام داد و پیشرفت‌های حاصل شده در زمینه هوش مصنوعی و پردازش زبان طبیعی را ارزیابی کرد.

۲.۴.۲ سنجش بر اساس نظر مدل داور

یک روش دیگر برای سنجش عملکرد یک مدل زبانی، استفاده از یک مدل زبانی دیگر به عنوان داور است. در این رویکرد، یک مدل زبانی مستقل به عنوان مرجع برای ارزیابی کیفیت پاسخ‌های تولید شده توسط مدل اصلی مورد استفاده قرار می‌گیرد. این روش به دلیل قابلیت‌های بالای مدل‌های زبانی در پردازش و درک زبان طبیعی، می‌تواند به طور موثری به ارزیابی دقت و کیفیت پاسخ‌ها کمک کند.

در این فرآیند، پاسخ‌های تولید شده توسط مدل اصلی به مدل قاضی ارائه می‌شود. مدل قاضی می‌تواند با استفاده از معیارهای مختلفی مانند شباهت معنایی با پاسخ اصلی^{۲۸}، صحت اطلاعات، و سازگاری با زمینه، کیفیت پاسخ‌ها را ارزیابی کند. به عنوان مثال، مدل قاضی می‌تواند با بررسی تطابق پاسخ‌ها با اطلاعات موجود در متون معتبر یا داده‌های آموزشی، نمره‌ای برای هر پاسخ تولید کرده [۱۲] یا پاسخی را بر پاسخ دیگر ترجیح دهد.

Massive Multitask Language Understanding^{۲۷}
ground truth^{۲۸}

۳.۴.۲ سنجش بر اساس استنتاج زبان طبیعی

با داشتن یک مجموعه داده مانند K-QA [۱۳] ، شامل پاسخ‌های تولید شده توسط انسان که به همراه توضیحات دقیقی دسته‌بندی شده‌اند، این پاسخ‌ها به عنوان "الرامی"^{۲۹} یا "مفید"^{۳۰} مشخص گشته‌اند. این دسته‌بندی نشان می‌دهد که آیا توضیحات باید به طور ضروری در پاسخ گنجانده شوند یا اینکه اضافی و مفید هستند.

این حقایق اتمی می‌توانند برای به کارگیری یک روش ارزیابی مبتنی بر استنتاج زبان طبیعی^{۳۱} استفاده شوند. در این روش، پاسخ مدل به عنوان "مقدمه"^{۳۲} و هر یک از توضیحات انسانی به عنوان "فرضیه"^{۳۳} در نظر گرفته می‌شود. سپس یک مدل زبانی توانا در حوزه استنتاج زبان طبیعی تعیین خواهد کرد که آیا مقدمه مستلزم^{۳۴}، متناقض^{۳۵} یا خنثی^{۳۶} با فرضیه است.

با انجام این کار روی همه رکوردهای مجموعه داده، دو امتیاز کامل بودن^{۳۷} و حقیقت داشتن^{۳۸} به صورت زیر به دست می‌آید.

$$S_{comp}(r_i, A'_i) = \frac{\lvert [r_i \quad entails \quad a] \rvert}{|A'_i|}$$

$$S_{fact}(r_i, A'_i) = \begin{cases} 0 & \text{if } \exists a \in A_i \quad s.t. \quad r_i \quad contradicts \quad a \\ 1 & \text{if } otherwise \end{cases}$$

۴.۴.۲ سنجش بر اساس امتیاز bert

در این معیار، کیفیت پاسخ‌های تولید شده توسط یک مدل زبانی فقط رمزگذار مانند bert ارزیابی می‌شود. [۱۴] در این فرآیند، پاسخ داده شده با پاسخ صحیح^{۳۹} موجود در مجموعه داده مقایسه می‌گردد.

must have ^{۴۰}
nice to have ^{۴۱}
natural language inference ^{۴۲}
premise ^{۴۳}
hypothesis ^{۴۴}
entailment ^{۴۵}
contradiction ^{۴۶}
neutral ^{۴۷}
completeness ^{۴۸}
factuality ^{۴۹}
ground truth ^{۵۰}

مدل زبانی پاسخ فقط رمزگذار را به دو بردار تبدیل می‌کند: یکی برای پاسخ داده شده و دیگری برای پاسخ صحیح. سپس برای محاسبه امتیاز bert، کافی است میزان شباهت این دو بردار را با یک روش شباهت‌سنجی مانند شباهت کسینوسی^{۴۰} محاسبه کنیم.

پس از محاسبه امتیازها برای تمامی پاسخ‌ها در مجموعه داده، این امتیازات جمع‌آوری و میانگین‌گیری می‌شوند. میانگین امتیازها نمای کلی از عملکرد مدل را ارائه می‌دهد و به شناسایی نقاط قوت و ضعف آن کمک می‌کند.

۵.۲ یادگیری تقویتی در مدل‌های زبانی

با داشتن یک مدل زبانی اولیه که بتواند برای هر پرسشی پاسخ خود را تولید کند، می‌توانیم یک مدل یا تابع پاداش^{۴۱} در نظر بگیریم و آن مدل زبانی را به‌گونه‌ای به‌روزرسانی کنیم که این تابع پاداش بیشینه گردد. به این صورت که ابتدا داده‌های ترجیحات انسانی (مانند مقایسه دو پاسخ و انتخاب بهتر) جمع‌آوری شده و با آن‌ها یک مدل پاداش آموزش داده می‌شود. سپس، با استفاده از الگوریتم‌های بهینه سازی که در ادامه به آن‌ها خواهیم پرداخت، سیاست تولید پاسخ مدل زبانی به تدریج اصلاح می‌گردد تا امتیاز پاداش بالاتری کسب کند. این شیوه باعث می‌شود که مدل زبانی نه تنها با ترجیحات انسانی هم راستا شود، بلکه پاسخ‌های مخرب، نادرست، جانبدارانه یا نامناسب کاهش یابد، ایمنی و قابلیت اطمینان افزایش یابد و کیفیت کلی مکالمه از نظر روانی بودن، مفید بودن و جذابیت به طور چشمگیری بهبود یابد. در نهایت، این روش به مدل‌های زبانی کمک می‌کند تا خروجی‌هایی تولید کنند که نه تنها دقیق و مرتبط باشند، بلکه با ارزش‌ها، فرهنگ و انتظارات کاربران انسانی نیز هماهنگ گردند.

۱.۵.۲ یادگیری تقویتی با بازخورد انسان یا هوش مصنوعی

برای مشخص کردن ترجیحات انسانی یاد شده علاوه بر انسان [۱۵]^{۴۲} می‌توان از بازخورد هوش مصنوعی نیز استفاده کرد؛ بدین صورت که از یک مدل هوش مصنوعی پیشرفته‌تر یا تخصصی (مانند یک مدل زبانی بزرگ‌تر یا یک سیستم ارزیابی خودکار) خواسته می‌شود که به پاسخی که یک مدل زبانی تولید کرده نمره‌ای عددی (مثلاً از ۱ تا ۱۰ بر اساس معیارهایی نظیر صحّت، مفید بودن، روانی و ایمنی) بدهد یا آن را در قیاس با پاسخ‌های دیگر تولیدشده توسط همان مدل یا مدل‌های متفاوت رتبه‌بندی کند. این روش که از آن با عنوان یادگیری تقویتی با

^{۴۰} cosine similarity
^{۴۱} reward model or function
^{۴۲} reinforcement learning with human feedback

با خورد هوش مصنوعی [۱۶]^{۴۳} یاد می‌شود، هزینه‌های جمع‌آوری داده‌های انسانی را به شدت کاهش می‌دهد، سرعت فرآیند را افزایش می‌دهد و امکان مقیاس‌پذیری گسترده‌تری فراهم می‌آورد، هرچند ممکن است به دلیل محدودیت‌های مدل ارزیابی‌کننده، برخی جنبه‌های ظریف فرنگی، اخلاقی یا زمینه‌ای انسانی را به طور کامل پوشش ندهد؛ بنابراین اغلب ترکیبی از بازخورد انسانی و هوش مصنوعی برای دستیابی به بهترین نتایج به کار گرفته می‌شود.

۲.۵.۲ الگوریتم‌های بهینه‌سازی مدل زبانی

الگوریتم‌های متعددی برای بهبود عملکرد مدل‌های زبانی بزرگ پیشنهاد شده‌اند که هر یک از رویکردهای متفاوتی در چارچوب یادگیری تقویتی^{۴۴} بهره می‌برند. از میان این روش‌ها، می‌توان به الگوریتم بهینه‌سازی سیاست مجاورتی^{۴۵} [۱۷] به عنوان یکی از پرکاربردترین و پایدارترین روش‌ها اشاره کرد. این الگوریتم با محدود کردن میزان تغییر در سیاست، موجب پایداری در فرآیند آموزش می‌شود. در این رویکرد، ابتدا یک مدل پاداش آموزش داده می‌شود و سپس مدل زبانی تلاش می‌کند تا با بیشینه‌سازی مقدار خروجی این مدل پاداش، عملکرد خود را بهبود دهد.

یکی دیگر از این الگوریتم‌ها، بهینه‌سازی سیاست بر پایه پاداش گروهی^{۴۶} [۱۸] است. این الگوریتم با معرفی پادash‌های نسبی درون‌گروهی، تلاش می‌کند مشکلات نوسان و ناپایداری موجود در الگوریتم بهینه‌سازی سیاست مجاورتی را کاهش دهد و از طریق ارزیابی نسبی پاسخ‌ها در یک گروه، بازخورد دقیق‌تر و پایدارتری برای آموزش فراهم کند.

در ادامه، الگوریتم بهینه‌سازی مستقیم ترجیحات^{۴۷} [۱۹] به جای تکیه بر یک مدل پاداش بیرونی، سعی دارد تابع هدف درونی^{۴۸} را در مدل زبانی بهینه کند. همانطور که در تصویر ۲.۲ می‌بینید در این روش، مدل به صورت مستقیم از جفت پاسخ‌های ترجیحی (پاسخ تاییدشده و پاسخ ردشده)^{۴۹} یاد می‌گیرد و نیازی به تخمين یا آموزش مدل پاداش مجزا ندارد. این ویژگی باعث سادگی و پایداری بیشتر در فرآیند هم‌ترازسازی مدل با ترجیحات انسانی می‌شود. تابع هزینه این الگوریتم به صورت زیر تعریف می‌شود

reinforcement learning with AI feedback^{۴۳}

Reinforcement Learning^{۴۴}

Proximal Policy Optimization (PPO)^{۴۵}

Group Reward Policy Optimization (GRPO)^{۴۶}

Direct Preference Optimization (DPO)^{۴۷}

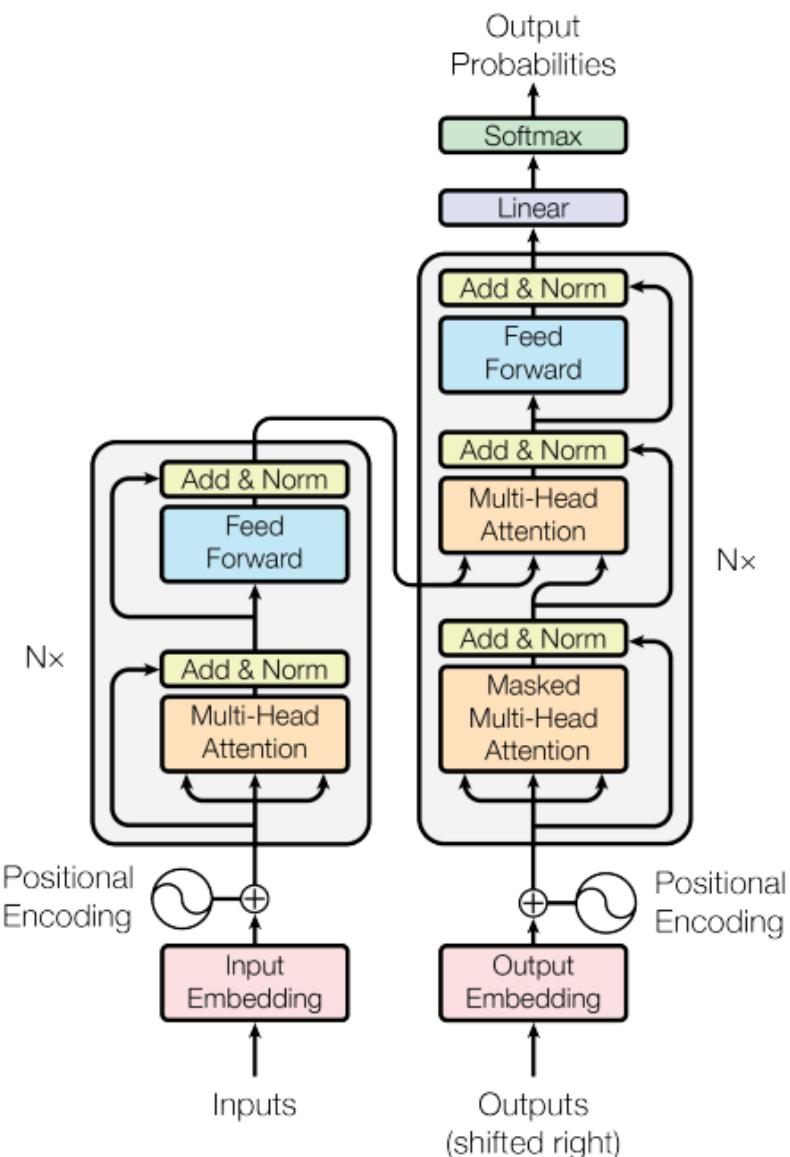
Intrinsic Objective Function^{۴۸}

Rejected–Preferred Answer Pairs^{۴۹}

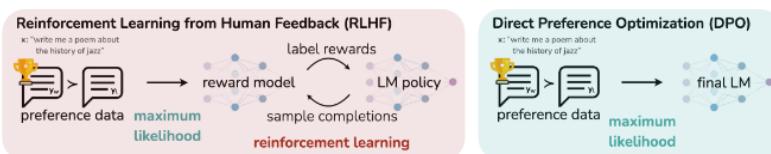
$$\mathcal{L}_{\text{DPO}}(\theta) = -\mathbb{E}_{(x,y^+,y^-)} \left[\log \sigma \left(\beta \left(\log \frac{\pi_\theta(y^+|x)}{\pi_{\text{ref}}(y^+|x)} - \log \frac{\pi_\theta(y^-|x)}{\pi_{\text{ref}}(y^-|x)} \right) \right) \right] \quad (1.2)$$

در نهایت، روش بهینه‌سازی مستقیم تابع کیو ^{۵۰} [۲۰] فرآیند تولید پاسخ را به صورت یک فرآیند تصمیم مارکوف ^{۵۱} مدل می‌کند و با بهره‌گیری از چارچوب بازیگر-منتقد نرم ^{۵۲} [۲۱] به طور مستقیم تابع Q را که به وسیله مدل زبانی پارامترگذاری شده است، بهینه می‌سازد. این مدل‌سازی بر پایه‌ی زنجیره تصمیم گیری مارکف، نسبت به روش‌های مبتنی بر Bandit از ساختار قوی‌تری برخوردار است و امکان نظارت فرایندی مؤثرتر را فراهم می‌کند. به طور کلی، هر یک از این الگوریتم‌ها با هدف افزایش کارایی، پایداری و هم‌خوانی رفتاری مدل‌های زبانی با ترجیحات انسانی توسعه یافته‌اند و بخش مهمی از پیشرفت‌های اخیر در حوزه‌ی آموزش و هم‌ترازسازی مدل‌های زبانی را تشکیل می‌دهند.

Direct Q-function Optimization (DQO)^{۵۰}
Markov Decision Process (MDP)^{۵۱}
Soft Actor-Critic (SAC)^{۵۲}



شكل ۱.۲: معماری ترنسفورمر



شكل ۲.۲: مقایسه الگوریتم بهینه‌سازی سیاست مجاورتی و بهینه‌سازی مستقیم ترجیحات

فصل ۳

بررسی کارهای پیشین

۱.۳ مقدمه

همان طور که پیشتر اشاره شد، علیرغم پیشرفت‌های چشمگیر در توسعه مدل‌های زبانی پزشکی به زبان انگلیسی، مانند توسعه و معرفی مدل‌های Med-Gemini [۲۴] یا مدل MedPalm [۲۳] [۲۲]، متأسفانه در حوزه زبان فارسی هنوز کار چندانی در این زمینه انجام نشده است. این مسئله بدین معناست که ما در حوزه زبان فارسی تقریباً با یک کاغذ سفید رو به رو هستیم. در این پایان‌نامه تلاش شده است تا قدمی رو به جلو در جهت توسعه مدل‌های زبانی پزشکی برای زبان فارسی برداشته شود. در ادامه، به بررسی کارهای پیشین انجام‌شده در حوزه زبان انگلیسی، حوزه زبان فارسی و مدل‌های دارای قابلیت دلیل‌آوری خواهیم پرداخت.

۲.۳ کارهای پیشین در حوزه زبان انگلیسی

۱.۲.۳ مدل‌های MedPaLM

مدل‌های MedPaLM یکی از مدل‌های زبانی پزشکی بزرگ^۱ است که توسط تیم تحقیقاتی گوگل برای کاربردهای پزشکی توسعه داده شده است. این مدل با استفاده از داده‌های تخصصی پزشکی و بالینی آموزش

large medical language models¹

دیده است. هدف اصلی این خانواده از مدل‌های زبانی پزشکی پاسخ‌گویی به پرسش‌های پزشکی با دقت بالا، کمک به پزشکان در تصمیم‌گیری‌های بالینی، و تسهیل دسترسی به اطلاعات پزشکی برای کاربران است. نسخه‌های مختلف این مدل، مانند MedPaLM و MedPaLM2 ، توانایی‌های قابل توجهی در درک و تحلیل زبان تخصصی پزشکی نشان داده‌اند و به عنوان یک ابزار نوین در حوزه هوش مصنوعی پزشکی شناخته می‌شوند. این مدل‌ها با استفاده از آزمون‌های استاندارد پزشکی (مانند USMLE) ارزیابی شده و توانسته‌اند عملکردی نزدیک به سطح متخصصین پزشکی ارائه دهند. مدل MedPaLM2 به عنوان یک گام مهم در جهت توسعه مدل‌های زبان تخصصی در حوزه سلامت و پزشکی شناخته می‌شود.

۲.۲.۳ مدل ChatDoctor

مدل ChatDoctor [۲۵] یکی از برجسته‌ترین تلاش‌ها در حوزه توسعه مدل‌های زبانی پزشکی است که شباهت قابل توجهی به فاز نخست پایان‌نامه حاضر دارد. تیم توسعه‌دهنده این مدل، داده‌های آموزشی خود را از دو پلتفرم آنلاین پرسش و پاسخ پزشکی به نام‌های iCliniq و HealthcareMagic جمع‌آوری کرده‌اند. این تیم ابتدا بیش از دویست هزار جفت پرسش و پاسخ پزشکی از این منابع گردآوری کرده و سپس با اعمال فیلترهایی بر اساس طول و کیفیت پاسخ‌ها، مجموعه‌ای با کیفیت بالا شامل صد هزار جفت پرسش و پاسخ نهایی ایجاد کرده‌اند. داده‌های مذکور به عنوان پایه‌ای برای آموزش و تنظیم دقیق ^۲ مدل LLaMa [۲۶] مورد استفاده قرار گرفته‌اند تا مدلی توانمند در تولید اطلاعات پزشکی دقیق و مرتبط ایجاد شود.

علاوه بر این، این مدل از رویکرد تولید مبتنی بر بازیابی اطلاعات ^۳ بهره برده است. این رویکرد به مدل امکان می‌دهد تا به اطلاعات جدید و خارجی دسترسی پیدا کرده و آن‌ها را به طور مؤثر در پاسخ‌های خود ادغام کند. چنین رویکردی موجب ارتقای عملکرد کلی سیستم شده و توانایی مدل در تولید پاسخ‌هایی دقیق‌تر و مرتبط‌تر را به طور چشمگیری بهبود بخشیده است.

۳.۲.۳ مدل‌های Meerkat

مدل‌های Meerkat [۲۷] یکی دیگر از تلاش‌های برجسته در حوزه توسعه مدل‌های زبانی پزشکی است. این پروژه با استخراج زنجیره‌های تفکر ^۴ از کتاب‌های درسی پزشکی و تنظیم دقیق یک مدل زبانی پایه با استفاده از این داده‌ها، همراه با مجموعه داده‌های مکمل دیگر، به وجود است. همانند فاز دوم پایان نامه حاضر هدف

^۲ fine-tuning
^۳ Retrieval-Augmented Generation (RAG)
^۴ chain of thought

اصلی Meerkat تمرکز بر فرآیندهای استدلالی است که در تصمیم‌گیری‌های پزشکی نقش دارند. این مدل تلاش کرده است تا نه تنها اطلاعات پزشکی دقیق ارائه دهد، بلکه فرآیندهای شناختی و تصمیم‌گیری متخصصان حوزه سلامت را شبیه‌سازی کند. به همین دلیل، Meerkat به عنوان مدلی برای تعاملات پیچیده‌تر و آگاهانه‌تر در حوزه پزشکی معرفی شده است.

۴.۲.۳ مدل MedMobile

[۲۸] تلاشی دیگر در حوزه مدل‌های زبانی کوچک پزشکی است. برای توسعه این مدل زبانی کوچک، مدل Phi-3-mini [۲۹] به عنوان مدل پایه^۵ استفاده از ترکیبی از داده‌های مصنوعی و تولیدشده توسط انسان تنظیم دقیق^۶ شده است تا عملکردی بهینه و مناسب برای اجرا روی دستگاه‌های همراه مانند موبایل ارائه دهد. با تمرکز بر نیازهای خاص کاربران دستگاه‌های همراه، MedMobile تلاش کرده است مدلی کارآمد و مؤثر فراهم کند که دسترسی به اطلاعات پزشکی باکیفیت را در هر زمان و مکان به صورت محلی^۷ ممکن می‌سازد.

۳.۳ کارهای پیشین در حوزه زبان فارسی

همان‌طور که پیش‌تر اشاره شد، تحقیقات محدودی بر روی مدل‌های زبانی پزشکی فارسی تمرکز داشته‌اند که این امر نشان‌دهنده شکاف قابل توجهی در منابع موجود برای جامعه پزشکی فارسی‌زبان است. علاوه بر این، پژوهش‌های بسیار اندک موجود در این زمینه، به طور کامل در مورد مجموعه داده‌ها، مدل‌ها و کدهای خود خود متن بسته^۸ هستند.

از سوی دیگر، تمامی این تلاش‌ها عمدتاً بر روی راهکارهای استخراجی^۹ متمرکز بوده‌اند که هدف‌شان بازیابی اطلاعات مرتبط از منابع از پیش تعریف شده است، به جای استفاده از رویکردهای تولیدی^{۱۰} که قادر به تولید پاسخ‌های آگاه از زمینه باشند.

baseline model ^۵
fine tune ^۶
local ^۷
closed-source ^۸
extractive ^۹
generative ^{۱۰}

۱.۳.۳ مدل Sina-bert

شاید اولین و برجسته ترین مدل زبانی پزشکی فارسی، Sina-BERT [۳۰] باشد که شامل آموزش یک مدل BERT [۳۱] با استفاده از یک پیکره خوش شده^{۱۱} همراه با مجموعه داده پرسش و پاسخ پزشکی فارسی است که به طور خاص برای کاربردهای مختلف از جمله پاسخ به سوالات پزشکی، تحلیل احساسات پزشکی و بازیابی سوالات پزشکی توسعه یافته‌اند.

Sina-BERT در میان تلاش‌های متمرکز بر زبان فارسی، بیشترین شباهت را به فاز نخست پایان نامه حاضر دارد؛ با این تفاوت که از مدل برتر^{۱۲} یک مدل زبانی مبتنی بر رمزگذار^{۱۳} به عنوان مدل پایه استفاده می‌کند. این انتخاب تولید پاسخ توسط این مدل را عمل ناممکن می‌سازد، چرا که برتر عمدتاً برای درک و استخراج اطلاعات طراحی شده است نه برای تولید پاسخ.

۲.۳.۳ سیستم پرسش و پاسخ پزشکی دکترویسی و همکاران

یکی از آثار برجسته در حوزه پردازش زبان طبیعی، سیستم پرسش و پاسخ پزشکی فارسی است که توسط دکتر ویسی و همکارانش [۳۲] طراحی و توسعه داده شده است. این سیستم به طور کلی شامل سه مأذول اصلی است: پردازش پرسش، بازیابی سند و استخراج پاسخ. مأذول پردازش پرسش وظیفه تحلیل و اصلاح پرسش‌های کاربران را بر عهده دارد تا پرسش‌ها به شکل بهینه برای مراحل بعدی آماده شوند. سپس، مأذول بازیابی سند با استفاده از الگوریتم‌های پیشرفته، اسناد پزشکی مرتبط را از میان داده‌های از پیش تعیین شده پیدا می‌کند. در نهایت، مأذول استخراج پاسخ با شناسایی دقیق اطلاعات موجود در اسناد بازیابی شده، مناسب‌ترین پاسخ‌ها را استخراج کرده و به کاربران ارائه می‌دهد. این سیستم نه تنها به طور مؤثر به پرسش‌های پزشکی پاسخ می‌دهد، بلکه ساختار مأذول‌ار آن امکان بهبود و توسعه در آینده را نیز فراهم می‌سازد.

۳.۳.۳ پایان نامه کارشناسی ارشد خانم لیلا دارابی

مشابه به این دو اثر، پیشین لیلا دارابی در پایان نامه ارشد خود [۳۳] از مدل‌هایی مانند Pars-BERT [۳۴] برای بازیابی پاسخ‌های مرتبط استفاده کرده است. رویکرد او شامل یافتن سوالات مشابه برای مدیریت پرسش‌های تکراری و به کارگیری استراتژی‌های ارزیابی دقیق و سهل‌گیرانه برای پاسخ‌های دقیق یا تقریبی می‌شود. علاوه بر

crawled^{۱۱}

BERT^{۱۲}

encoder-based^{۱۳}

این، روش‌های طبقه‌بندی و شناسایی موجودیت‌های نامدار^{۱۴} برای بهبود ارتباط پاسخ‌ها از طریق دسته‌بندی سوالات و شناسایی موجودیت‌های پزشکی مانند نام داروها و بیماری‌ها به کار گرفته می‌شوند.

۴.۳ بهبود قابلیت دلیل آوری در حوزه پزشکی

۱.۴.۳ مدل MedSSS

یکی از تلاش‌های برجسته در این حوزه، مدل زبانی MedSSS است. [۳۵] هدف اصلی این چارچوب، ارتقای توانایی استدلال مدل‌های پزشکی از طریق دقیق‌تر بخشی و ریزسازی مراحل میانی استدلال بین پرسش و پاسخ نهایی است. همانطوری که در تصویر ۱.۳ برای دستیابی به این هدف، پژوهشگران از الگوریتم جستجوی درختی مونت‌کارلو^{۱۵} [۳۶] استفاده کردند. این الگوریتم بر روی مجموعه‌داده‌هایی از پرسش‌های چندگزینه‌ای پزشکی به کار گرفته شد تا مسیرهای استدلالی ساختارمند و مرحله‌به‌مرحله تولید شود؛ مسیرهایی که نشان می‌دهند مدل چگونه از پرسش به پاسخ درست می‌رسد. بر اساس این رویکرد، نویسنده‌گان به صورت مصنوعی سه مجموعه‌داده مجزا ساختند که هر یک در مرحله‌ای خاص از آموزش مدل نقش دارد:

۱. مجموعه‌داده تنظیم دقیق نظارت شده: برای آموزش مدل سیاستی^{۱۶} از طریق نمونه‌های استدلال باکیفیت.

۲. مجموعه‌داده پاسخ‌های ترجیح‌داده شده و رشدده: برای آموزش مبتنی بر بهینه‌سازی مستقیم ترجیحات^{۱۷} تا مدل بتواند میان پاسخ‌های بهتر و ضعیف‌تر تمایز قائل شود.

۳. مجموعه‌داده برچسب‌های نرم دوسویه: جهت تنظیم دقیق مدل پاداش فرآیند^{۱۸} که به جای تمرکز صرف بر پاسخ نهایی، کیفیت فرآیند استدلال مدل را ارزیابی می‌کند.

در نهایت، مدل سیاستی آموزش دیده به عنوان موتور اصلی استدلال در فرآیند تولید پاسخ عمل می‌کند، در حالی که مدل پاداش فرآیند آموزش دیده نقش راهنمای ارزیاب را در هنگام تولید ایفا می‌نماید. این مدل دوم کمک می‌کند تا فرآیند استدلال مدل به صورت پیوسته اصلاح شده و از نظر منطق، انسجام و درستی پزشکی ارزیابی گردد.

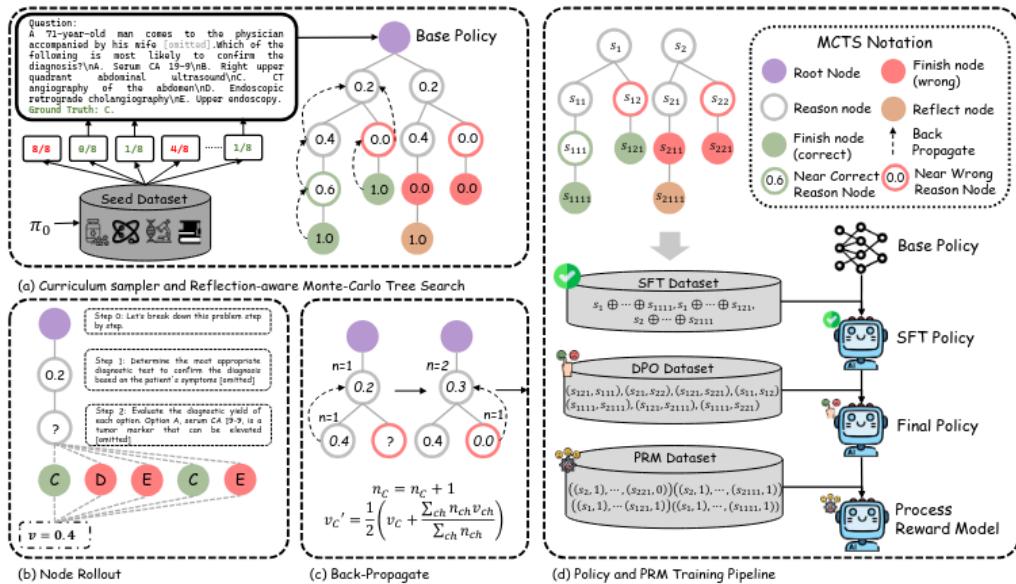
^{۱۴} (NER) Recognition Entity Named

^{۱۵} monte carlo tree search

^{۱۶} policy model

^{۱۷} direct preference optimization

^{۱۸} process reward model



شکل ۱۰.۳: بلوک دیاگرام مدل medsss

۲.۴.۳ مدل MedReason

یکی دیگر از دستاوردهای مهم در زمینه توسعه مدل‌های زبانی استدلال محور در حوزه پزشکی، مدل زبانی MedReason است. [۳۷] در این پژوهش، نویسندهای از یک گراف دانش ساختاریافته پزشکی^{۱۹} بهره گرفتند تا جفت‌های متدالول پرسش و پاسخ را به مسیرهای استدلالی دقیق و گام‌به‌گام تبدیل کنند. هر مسیر استدلالی، زنجیرهای منطقی از گام‌های مرتبط را نشان می‌داد که یک پرسش بالینی را به پاسخ صحیح آن متصل می‌کرد و این ارتباط بر پایه روابط دانش پزشکی مانند علائم، تشخیص‌ها، درمان‌ها و سازوکارهای فیزیولوژیکی بناده بود. پژوهشگران با ساخت چنین مجموعه‌دادهای غنی شده از استدلال، توانستند یک مدل زبانی پایه را بر اساس این نمونه‌های ساختاریافته تنظیم دقیق کنند. این رویکرد سبب شد تا توانایی مدل در انجام وظایف پیچیده و استدلال محور در حوزه پزشکی به طور قابل توجهی بهبود یابد. نتایج حاصل نشان داد که نظرات مبتنی بر استدلال استخراج شده از گراف‌های دانش^{۲۰}، می‌توانند نقش مؤثری در ارتقای دقت و قابلیت تحلیل مدل‌های زبانی پزشکی ایفا کند.

یافته‌های حاصل از پژوهش MedReason همچنین بر اهمیت نمایش‌های ساختاریافته استدلال در بهبود تفسیرپذیری^{۲۱} و عمق تحلیلی مدل‌های زبانی پزشکی تأکید می‌کنند. این چارچوب نشان می‌دهد که ترکیب دانش ساختاریافته با فرآیند یادگیری زبانی، می‌تواند پلی میان درک داده‌محور و استدلال مبتنی بر دانش در سامانه‌های

Medical Knowledge Graph^{۱۹}
Knowledge Graph-based Reasoning Supervision^{۲۰}
Interpretability^{۲۱}

هوش مصنوعی پزشکی ایجاد کند.

۳.۴.۳ مدل HuatuoGPT-01

یکی از پیشرفت‌های مهم دیگر در توسعه مدل‌های زبانی استدلال محور پزشکی، چارچوبی با نام-HuatuoGPT-01 است. [۳۸] در این پژوهش، نویسندهان یک چارچوب استدلال با هدایت اعتبارسنجی^{۲۲} را معرفی کردند که هدف آن بهبود انسجام منطقی و دقت مسیرهای استدلالی تولیدشده توسط مدل است.

در این رویکرد، یک مدل اعتبارسنج^{۲۳} به عنوان جز کنترلی و ارزیاب به کار گرفته شد تا در حین فرآیند تولید استدلال، عملکرد مدل سیاستی را ارزیابی و هدایت کند. این مکانیزم اطمینان حاصل می‌کرد که هر مسیر استدلالی تولیدشده با واقعیت‌های علمی و اصول پذیرفته شده پزشکی سازگار باشد. با استفاده از فرآیند فیلترسازی و پالایش مسیرهای استدلال از طریق این مدل اعتبارسنج، پژوهشگران موفق به تولید یک مجموعه‌داده باکیفیت متشکل از دنباله‌های استدلالی تأییدشده شدند.

در گام بعد، نویسندهان از ترکیب دور رویکرد آموزش، شامل تنظیم دقیق نظرارت شده و یادگیری تقویتی، برای آموزش مدل زبانی پایه با استفاده از داده‌های اعتبارسنجی شده بهره بردن. این رویکرد دوگانه باعث شد مدل نه تنها رفتارهای استدلالی صحیح را تقلید کند، بلکه اصول استدلال منطقی را از طریق بهینه‌سازی مبتنی بر پاداش درونی‌سازی نماید. نتایج حاصل از این پژوهش نشان داد که مدل HuatuoGPT-01 در طیف وسیعی از وظایف پزشکی، از جمله پاسخ‌گویی به پرسش‌های بالینی و استدلال تشخیصی، بهبود چشمگیری در دقت و قابلیت اعتماد از خود نشان می‌دهد. این دستاورد اهمیت چارچوب‌های یادگیری هدایت شده با اعتبارسنج را در ارتقای مدل‌های زبانی استدلالی در حوزه پزشکی برجسته می‌سازد و مسیر توسعه سامانه‌های هوش مصنوعی با استدلال قابل اعتماد را هموار می‌کند.

۵.۳ بهبود قابلیت دلیل آوری در حوزه‌های دیگر

۱.۰.۵.۳ مدل Deepseek

یکی از تاثیرگذارترین پژوهش‌های اخیر در حوزه گسترده استدلال در هوش مصنوعی، مدل DeepSeek-R است. [۳۹] این مدل بر پایه نسخه پایه DeepSeek-V3-Base توسعه یافته و با هدف ارتقای صریح قابلیت‌های

Verification-Guided Reasoning Framework^{۲۲}
Verifier Model^{۲۳}

استدلالی، چارچوبی مبتنی بر یادگیری تقویتی را به کار می‌گیرد. در این چارچوب، نویسنده‌گان از الگوریتم بهینه سازی سیاست بر پایه پاداش گروهی^{۲۴} استفاده کردند تا فرآیند یادگیری تقویتی را به گونه‌ای هدایت کنند که عملکرد استدلالی مدل به صورت هدفمند بهبود یابد.

در این رویکرد، تابع پاداش به طور ویژه طراحی شد تا کیفیت استدلال مدل را ارزیابی کرده و آن را به حداکثر برساند. نتایج نشان داد که مدل R از طریق این الگوی آموزشی، پیشرفت چشمگیری در منطق استدلالی و دقت حل مسئله در مجموعه‌ای از معیارهای استاندارد^{۲۵} به دست آورده است. با این حال، استفاده از یادگیری تقویتی بدون کنترل کامل، پیامدهایی جانبی نیز به همراه داشت؛ از جمله کاهش روانی زبان، افت انسجام جمله‌ای، و افزایش بروز پدیده اختلاط زبانی^{۲۶} در خروجی مدل.

برای رفع این چالش‌ها، پژوهشگران مرحله‌ای تکمیلی به فرآیند آموزش افزودند. در این مرحله، مقدار اندکی داده نظارت شده اولیه^{۲۷} مورد استفاده قرار گرفت و یک زنجیره آموزشی چندمرحله‌ای^{۲۸} طراحی شد. این فاز تنظیم دقیق کمک کرد تا ضمن حفظ توانمندی‌های استدلالی کسب شده از یادگیری تقویتی، روانی و خوانایی زبان طبیعی نیز بازگردانده شود.

در نتیجه، مدل R DeepSeek-R به عنوان گامی اساسی در مسیر توسعه هوش مصنوعی استدلال محور مطرح شد؛ مدلی که نشان می‌دهد یادگیری تقویتی می‌تواند به طور قابل توجهی توانایی استدلال را ارتقا دهد، مشروط بر آنکه با تنظیم دقیق و نظارت زبانی هدفمند ترکیب شود تا تعادل میان کیفیت زبانی و عمق استدلالی حفظ گردد.

۲.۵.۳ چهارچوب Thought Preference Optimization

یکی از دستاوردهای قابل توجه در زمینه بهبود توانایی استدلال در مدل‌های زبانی، روش بهینه سازی ترجیحات فکری^{۲۹} است. [۴۰] در این پژوهش، نویسنده‌گان چارچوبی مبتنی بر ترجیحات را برای ارتقای کیفیت استدلال در مدل‌های زبانی پیشنهاد کردند.

همانطور که در تصویر ۲.۳ می‌بینید در این رویکرد، هنگامی که یک پرسش به مدل ارائه می‌شود، ابتدا مدل چندین مسیر استدلالی متفاوت را به عنوان پاسخ‌های بالقوه تولید می‌کند. سپس این مسیرهای استدلالی توسط یک مدل داور^{۳۰} مورد ارزیابی قرار می‌گیرند. مدل داور با بررسی صحت منطقی، انسجام و کیفیت استدلال،

group reward policy optimization^{۲۴}

Benchmarks^{۲۵}

Language Mixing^{۲۶}

Cold-start Supervised Data^{۲۷}

Multi-stage Training Pipeline^{۲۸}

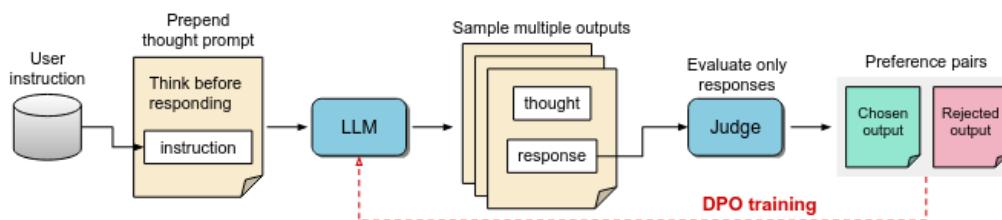
Thought Preference Optimization (TPO)^{۲۹}

Judge Model^{۳۰}

بهترین و ضعیف‌ترین نمونه‌ها را از میان خروجی‌ها شناسایی می‌کند.

در گام بعد، جفت‌های انتخاب‌شده از پاسخ‌های «بهتر» و «ضعیف‌تر» برای آموزش مدل پایه با استفاده از روش بهینه‌سازی مستقیم ترجیحات^{۳۱} به کار گرفته می‌شوند. این فرآیند باعث می‌شود مدل به صورت تدریجی مسیرهای استدلالی با کیفیت‌تر را ترجیح دهد و یاد بگیرد که استدلال‌هایی منطقی‌تر، پیوسته‌تر و نزدیک‌تر به شیوه تفکر انسانی تولید کند.

نتایج این پژوهش نشان داد که به کارگیری چارچوب TPO موجب بهبود چشمگیری در کیفیت استدلال مدل‌های زبانی می‌شود. این یافته بیانگر آن است که ارتقای توانایی استدلالی مدل‌ها تنها از طریق افزایش مقیاس یا داده‌های آموزشی نظارت‌شده حاصل نمی‌شود، بلکه می‌توان با بهره‌گیری از سازوکارهای یادگیری مبتنی بر ترجیح و بازخورد ارزیابانه، به بهبودی مؤثرتر و پایدارتر در فرآیند استدلال مدل دست یافت.



شکل ۲.۳: بلوک دیاگرام Thought Preference Optimization

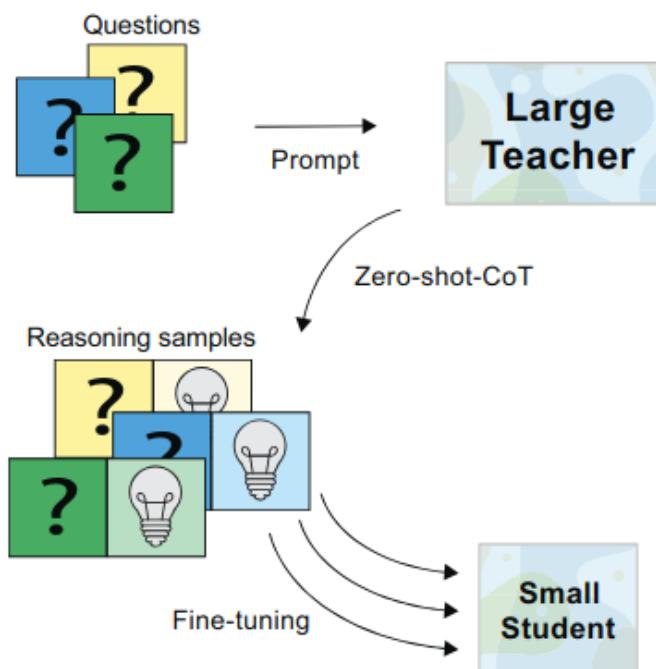
۳.۵.۳ آموزش استدلال مدل‌های بزرگ به مدل‌های کوچک

یکی از پژوهش‌های مرتبط در زمینه انتقال توانایی استدلال میان مدل‌های زبانی، توسط آقای هو و همکاران انجام شده است.^{۴۱} در این پژوهش، نویسنده‌گان به بررسی چگونگی انتقال قابلیت‌های استدلالی از مدل‌های زبانی بزرگ به مدل‌های کوچک‌تر پرداختند. ایده اصلی این رویکرد بر آن استوار است که مدل‌های کوچک‌تر می‌توانند از داده‌های تولیدشده توسط مدل‌های بزرگ‌تر، که از قدرت استدلالی بالاتری برخوردارند، برای بهبود عملکرد خود بهره ببرند.

همانطور که در تصویر ۳.۳ در این روش، ابتدا یک مدل زبانی بزرگ با عملکرد استدلالی قوی‌تر به تولید مسیرهای استدلالی^{۳۲} و جفت‌های پرسش و پاسخ می‌پردازد. این خروجی‌ها به عنوان داده‌های آموزشی با کیفیت و مبتنی بر استدلال، برای تنظیم دقیق مدل کوچک‌تر مورد استفاده قرار می‌گیرند. بدین ترتیب، مدل کوچک‌تر از این طریق، الگوهای حل مسئله و استراتژی‌های استدلالی را از مدل بزرگ‌تر فرا می‌گیرد.

Direct Preference Optimization(DPO)^{۳۱}
Reasoning Trajectories^{۳۲}

نتایج نشان داد که مدل کوچکتر پس از این آموزش هدفمند، توانست عملکردی رقابتی در وظایف استدلالی از خود نشان دهد؛ در حالی که هزینه محاسباتی آن به طور قابل توجهی کمتر از مدل بزرگتر بود. این پژوهش اهمیت استفاده از داده‌های مصنوعی مبتنی بر استدلال و یادگیری انتقالی هدفمند را در بهبود بازده و کارایی مدل‌های زبانی کوچکتر برجسته می‌کند. به عبارت دیگر، توانایی استدلال را می‌توان به صورت مؤثر میان مدل‌هایی با مقیاس‌های متفاوت منتقل کرد، مشروط بر آنکه فرآیند تنظیم دقیق بر اساس داده‌های تولیدشده از مدل‌های استدلالی پیشرفته انجام گیرد.



شکل ۳.۳: بلوک دیاگرام آموزش دلیل آوری توسط مدل بزرگتر به مدل کوچکتر

فصل ۴

جمع آوری دادگان

۱.۴ مقدمه

همان طور که پیشتر اشاره شد، در حوزه زبان فارسی نه مدل‌های عمومی موجود هستند و نه مجموعه داده‌های مناسب برای استفاده در پژوهش‌های مرتبط. بنابراین، برای پیشبرد این پایان‌نامه، ناچار به جمع‌آوری دادگان اختصاصی بودیم تا بتوانیم نیازهای تحقیقاتی را برآورده کنیم. فرآیند جمع‌آوری دادگان شامل روش‌هایی مانند ترجمه^۱ داده‌های موجود از زبان‌های دیگر و خزش داده‌ها از منابع مختلف برای ایجاد یک مجموعه داده جامع و کاربردی بوده است.

۲.۴ معرفی پیکره پزشکی فارسی

عدم وجود یک پیکره پزشکی اختصاصی به زبان فارسی، چالشی قابل توجه برای پژوهشگران و توسعه‌دهندگانی ایجاد می‌کند که هدف‌شان توسعه مدل‌های پزشکی در زبان فارسی است. بدون داده‌های متى باکیفیت و تخصصی که برای آموزش مدل‌های هوش مصنوعی ضروری است، این تلاش‌ها ممکن است با موانع رو به رو شوند و در نهایت بر توسعه فناوری‌ها و راه حل‌های پیشرفته پزشکی مناسب برای جمعیت فارسی‌زبان تاثیر بگذارند. برای حل این مشکل، ما یک مجموعه داده جامع شامل تقریباً نود میلیون توکن و حدود صد هزار مقاله گردآوری کردایم.^۲

^۱ ترجمه می‌تواند به صورت ماشینی یا انسانی انجام شود.

^۲ برای بازدید از این پیکره میتوانید به آدرس huggingface.co/datasets/gaokerena/medical_corpus مراجعه کنید.

گارسیا فررو و همکاران [۴۲] مجموعه‌ای از متون پزشکی را که به چهار زبان (انگلیسی، فرانسوی، اسپانیایی و ایتالیایی) اختصاص داشت، گردآوری کردند که می‌توان آن را همانطور که در جدول ۱.۴ نشان داده شده است با مجموعه ما مقایسه کرد. پیکره‌ای که ما گردآوری کرد های آنلاین پزشکی خوش شده است که میتوانید سهم هر مجله در این پیکره را در تصویر ۱.۴ ببینید.

گردآورنده	تعداد پرسش و پاسخ‌ها	زبان
I. Garcia Ferrero et al.	1.1B	انگلیسی
I. Garcia Ferrero et al.	950M	اسپانیایی
I. Garcia Ferrero et al.	675M	فرانسوی
I. Garcia Ferrero et al.	143M	ایتالیایی
ما	90M	فارسی

جدول ۱.۴: مقایسه پیکره گردآوری شده با پیکره‌های گردآوری شده توسط I. Garcia Ferrero et al.

۳.۴ معرفی مجموعه داده MF3QA

گردآوری یک مجموعه داده واقعی از پرسش و پاسخ‌های پزشک و بیمار اهمیت بسیاری در ارتقا توانایی‌های مدل‌های زبانی در حوزه بهداشت و درمان دارد. چنین مجموعه داده‌ای به مدل‌ها امکان می‌دهد تا اطلاعات ارزشمندی را که از تعاملات واقعی میان ارائه‌دهنگان خدمات بهداشتی و بیماران به دست می‌آید، بیاموزند. با تحلیل این تعاملات واقعی، مدل‌های زبانی می‌توانند به درک جزئیات اصطلاحات پزشکی، نگرانی‌های بیماران، و زمینه پیرامون سوالات بهداشتی دست یابند. علاوه بر این، این مجموعه داده مدل‌ها را قادر می‌سازد نه تنها محتوای دقیق پاسخ‌ها، بلکه ساختار و لحن مناسب برای پاسخ‌دهی به سوالات را نیز یاد بگیرند. این فرآیند دوگانه یادگیری از اهمیت بالایی برخوردار است، زیرا به مدل امکان می‌دهد پاسخ‌هایی دقیق، همدلانه و متناسب با زمینه ارائه دهد و در نهایت ارتباط و پشتیبانی از بیماران در محیط‌های پزشکی را بهبود بخشد.

در این زمینه، یانگ لیو در مقاله موری ^۳ خود [۴۳] به چندین مجموعه داده واقعی پرسش و پاسخ پزشک و بیمار اشاره کرده است، مقایسه‌ای میان این مجموعه دادگان و مجموعه داده ما در جدول ۲.۴ ارائه شده است.

^۳ survey

نام مجموعه داده	زبان	تعداد پرسش و پاسخ ها	گردآورنده
ChatDoctor	انگلیسی	100K	[۲۵] Yunxiang Li et al.
CMTMedQA	چینی	68K	[۴۴] Songhua Yang et al.
DISC-Med-SFT	چینی	465K	[۴۵] Zhijie Bao et al.
HuatuoGPT-sft-data-v1	چینی	226K	[۴۶] Hongbo Zhang et al.
Huatuo-26M	چینی	26M	[۴۷] Jianquan Li et al.
MedDialog	چینی و انگلیسی	3.66M	[۴۸] Guangtao Zeng et al.
Medical-Meadow	انگلیسی	160k	[۴۹] Tianyu Han et al.
MF3QA	فارسی	20k	ما

جدول ۲.۴: مقایسه مجموعه داده های پرسش و پاسخ آزاد پزشکی با مجموعه داده گردآوری شده

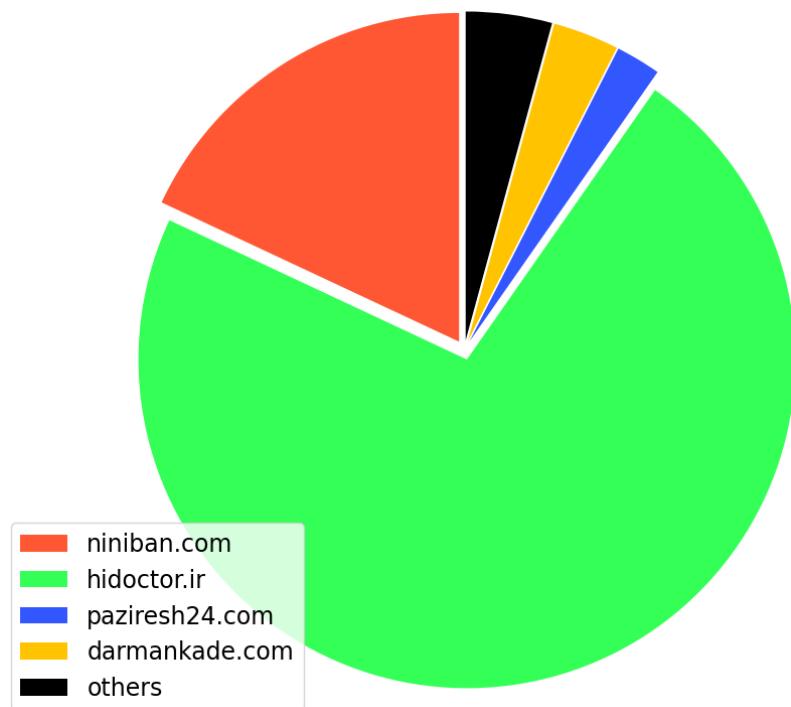
۱.۳.۴ منابع مجموعه داده MF3QA

همان طور که در شکل ۲.۴ نشان داده شده است، برای گردآوری مجموعه داده MF3QA مراحل مختلفی طی شده است. در بخش آموزش، پرسش و پاسخ های بیمار و پزشک موجود در تالارهای گفت و گویی پزشکی فارسی^۴ "دکترهست" و "نی نی بان" را خوش کرده ایم. برای بخش اعتبارسنجی، تنها از داده های موجود در سایت "نی نی بان" استفاده کرده ایم تا انسجام بیشتری در این بخش حاصل شود. در بخش آزمایش نیز، از سایت های "دکتریاب" و "ایزو ویزیت" بهره برده ایم و به منظور اطمینان از تنواع داده ها، مجموعه داده پرسش و پاسخ K-QA [۱۳] را ترجمه کرده و به این بخش اضافه کرده ایم.

۲.۳.۴ فیلتر کردن رکوردهای مجموعه داده MF3QA

در پایان نامه حاضر، بیش از صد و هشتاد هزار جفت پرسش و پاسخ از تالارهای گفت و گویی پزشکی فارسی گردآوری شده است. این جفت های پرسش و پاسخ، چه به صورت دستی^۵ و چه به صورت خودکار، مورد بررسی

^۴ Persian medical forums^۵ فرآیند فیلتر کردن دستی توسط خانم زهرا کاظمی و آقای میلاد توکلی، از دانشجویان کارشناسی مهندسی کامپیوتر، انجام شده است.



شکل ۱.۴: سهم هر مجله در پیکره پزشکی فارسی گردآوری شده

قرار گرفته و جفت‌هایی که حاوی اطلاعات مفید نبودند، حذف شده‌اند.^۶

این رویکرد مشابه کاری است که یونشیانگ لی و همکارانش برای توسعه مدل زبانی Chat Doctor انجام داده‌اند. [۲۵] آنها نیز داده‌ها را از تالارهای گفت‌وگوی پزشکی انگلیسی استخراج کرده و نیمی از جفت‌های پرسش‌وپاسخ را بر اساس طول پاسخ‌ها کنار گذاشته‌اند^۷، چراکه پاسخ‌های کوتاه‌تر معمولاً حاوی اطلاعات مفیدی نیستند. با این حال، ما با چالش بزرگ‌تری مواجه بودیم؛ پزشکان فارسی‌زبان معمولاً پاسخ‌های بسیار کوتاه‌تری نسبت به همتایان انگلیسی خود ارائه می‌دهند. این امر ما را مجبور کرد تا بیش از هشتاد درصد از رکوردهای پرسش‌وپاسخ خود را برای تضمین کیفیت کنار بگذاریم.

^۶ برای بازدید از مجموعه داده MF3QA به آدرس huggingface.co/datasets/gaokerena/MF3QA و برای بازدید از صد و هشتاد هزار جفت پرسش‌وپاسخ خوش شده به آدرس huggingface.co/datasets/gaokerena/MF3QA_uncleaned مراجعه کنید.

^۷ فیلتر کردن آنها صرفاً بر اساس طول پاسخ بوده ولی همانطور که پیشتر اشاره شد ما برای فیلتر کردن از روش‌های دستی نیز استفاده کرده‌ایم.



شکل ۲.۴: سهم هر تالار گفتگو در مجموعه داده MF3QA

۱.۲.۳.۴ خوش از تالار گفتگو دکترهست

خوش از تالار گفتگوی "دکترهست"، که اصلی‌ترین منبع مجموعه داده MF3QA است، با چالش خاصی همراه بود. این تالار گفتگو تمام رکوردهای تعامل پزشک و بیمار خود را به صورت مستقیم در سایت ارائه نمی‌دهد و فقط به دو هزار رکورد آخر دسترسی می‌دهد. علاوه بر این، هر رکورد به صد رکورد مرتبط دیگر پیوند داده شده است.

برای حل این چالش، از الگوریتم ۱.۴ استفاده شد. در این روش، داده‌های تالار گفتگو به صورت یک گراف در نظر گرفته شده و با استفاده از جستجوی عرض-اول^۸ توانستیم حدود صد و بیست هزار رکورد از مجموع دویست هزار رکورد موجود در این تالار گفتگو را استخراج کنیم. این فرایند حدود دو هفته طول کشید.

۴.۴ ترجمه قسمت پزشکی مجموعه داده MMLU

مجموعه داده MMLU [۱۱]^۹ یکی از معتربرترین مجموعه داده‌ها برای ارزیابی توانایی مدل‌های زبانی در درک و پاسخ‌دهی به سوالات چندوظیفه‌ای است. این مجموعه شامل سوالاتی در موضوعات مختلف از جمله علوم پزشکی، مهندسی، علوم انسانی و دیگر حوزه‌ها است که به صورت چندگرینه‌ای طراحی شده‌اند. در پژوهه ما، برای ارزیابی مدل زبانی پزشکی توسعه یافته، بخش پزشکی این مجموعه داده را به زبان فارسی ترجمه کردیم.^{۱۰} هدف از این کار، تطبیق داده‌های ارزیابی با زبان مورد استفاده در مدل و بررسی توانایی مدل در پاسخ‌دهی دقیق به سوالات تخصصی پزشکی در زبان فارسی بود.^{۱۱}

⁸ breadth first search^۸

⁹ Massive Multitask Language Understanding^۹

^{۱۰} ترجمه توسط آقای امیرحسین پورسینا دانشجوی پزشکی انجام شده است.

^{۱۱} برای بازدید از ترجمه این مجموعه داده به آدرس huggingface.co/datasets/gaokerena/FA_MED_MMLU مراجعه کنید.

الگوریتم ۱۰.۴ جستجو اول عرض برای استخراج رکورد های پرسش و پاسخ پزشکی

ورودی: گره های دارای دسترسی در تالار گفتگو (برگ ها)

خروجی: مجموعه ای از گره های بازدید شده

۱: یک پشته خالی S ایجاد کن

۲: یک مجموعه خالی $Visited$ ایجاد کن

۳: گره مبدأ v را به پشته S اضافه کن

۴: تا زمانی که پشته S خالی نیست انجام بده

۵: یک گره u را از پشته S بردار

۶: اگر گره u بازدید نشده است آنگاه

۷: گره u را به مجموعه $Visited$ اضافه کن

۸: برای هر همسایه n از گره u انجام بده

۹: اگر گره n بازدید نشده است آنگاه

۱۰: گره n را به پشته S اضافه کن

۱۱: پایان شرط اگر

۱۲: پایان حلقه برای

۱۳: پایان شرط اگر

۱۴: پایان حلقه تا زمانی که

۱۵: بازگردان نود های بازدید شده

۵.۴ گردآوری سوالات کنکور علوم پایه پزشکی ایران

آزمون علوم پایه پزشکی یک آزمون سراسری در ایران است که دانشجویان پزشکی موظف هستند پس از گذراندن دروس علوم پایه معمولاً در پنج ترم در آن شرکت کنند. این آزمون به منظور سنجش میزان آموخته های دانشجویان از دروس علوم پایه و آمادگی آنها برای ورود به مراحل بالینی برگزار می شود. در صورتی که دانشجو پس از سه مرتبه در این آزمون قبول نشود^{۱۲} از ادامه تحصیل در رشته پزشکی محروم می شود.

برای سنجش دانش پزشکی مدل زبانی خود ما سوالات این آزمون را از pdf سوالاتی که سازمان سنجش برای آن منتشر می کند استخراج کرده ایم.^{۱۳}^{۱۴}

^{۱۲} نمره قبولی در این آزمون در سالهای مختلف متفاوت است و معمولاً حدود سی و شش درصد می باشد

^{۱۳} برای این کار از کتابخانه fitz پایتون استفاده شده است.

^{۱۴} برای بازدید از این مجموعه داده به آدرس <https://huggingface.co/datasets/gaokerena/KOPP> مراجعه کنید.

٦.٤ ترجمه ماشینی مجموعه داده MedMCQA

برای پیاده‌سازی روش پیشنهادی خود، نیاز به یک مجموعه داده‌ی پرسش‌های چندگزینه‌ای پزشکی به زبان فارسی داشتیم. از آنجا که در زمان انجام پژوهش، چنین مجموعه داده‌ای در دسترس نبود، تصمیم گرفتیم از مدل زبانی بزرگ DeepSeek-V3 [۵۰] که از نظر هزینه و کارایی مقرن‌به‌صرفه است، برای ترجمه‌ی بخشی از مجموعه داده MedMCQA [۵۱] از زبان انگلیسی به فارسی استفاده کنیم.

به منظور حفظ نوع موضوعی، پرسش‌ها به صورت تصادفی از مجموعه داده MedMCQA انتخاب شدند تا دامنه‌ی وسیعی از مباحث علوم پزشکی — از جمله فیزیولوژی، داروشناسی، پاتولوژی، و سایر حوزه‌ها — در مجموعه‌ی نهایی پوشش داده شود.

برای اطمینان از کیفیت ترجمه‌ها، از دو داور زبانی مستقل به نام‌های grok-3-mini [۵۲] و gpt-4.1-mini [۵۳] استفاده کردیم. هر یک از این داوران موظف بودند که ترجمه‌ی هر پرسش و گزینه‌های آن را از نظر دقت معنایی، روانی زبانی و حفظ اصطلاحات تخصصی ارزیابی کنند. ترجمه تنها در صورتی به عنوان «تاییدشده» در نظر گرفته می‌شد که هر دو داور به آن امتیاز کامل پنج از پنج اختصاص دهند.

در پایان این فرآیند چند مرحله‌ای، موفق شدیم حدود ۱۸۰۰۰ پرسش چندگزینه‌ای پزشکی به زبان فارسی را که از نظر کیفیت ترجمه تایید شده بودند، گردآوری کنیم.

٧.٤ مجموعه دادگان PersianMedQA و PerMedCQA

پس از اتمام فاز جمع آوری داده‌ها و انتشار آن‌ها به صورت عمومی، دو مقاله مهم در حوزه جمع آوری مجموعه داده‌های پزشکی منتشر شد. مقاله نخست به معرفی مجموعه داده PerMedCQA [۵۴] پرداخته و مقاله دوم به معرفی مجموعه داده PersianMedQA [۵۵] اختصاص داشت.

مجموعه داده PerMedCQA از نظر ساختار و نوع پرسش‌ها شباهت زیادی به مجموعه داده MF3QA دارد. نویسنده مقاله PerMedCQA ادعا کرده است که این مجموعه داده نخستین مجموعه داده در نوع خود است. با این حال، بررسی شماره آرشیو این مقاله نشان می‌دهد که این ادعا نادرست است، زیرا مجموعه داده MF3QA پیش از PerMedCQA معرفی شده بود.

برای گسترش مجموعه داده خود و افزایش تنوع منابع، ما از مجموعه داده PersianMedQA استفاده کردیم. در این راستا، ۱۰۰۰ پرسش به صورت تصادفی از بخش آموزشی این مجموعه داده انتخاب و به مجموعه داده اصلی ما اضافه شد. این اقدام موجب افزایش داده‌ها و تنوع منابع مجموعه داده شد.

فصل ۴: جمع آوری دادگان

با اضافه کردن این پرسش‌ها به ۱۸۰۰۰ پرسشی که از ترجمه ماشینی مجموعه‌داده MedMCQA به دست آمده بود، در نهایت مجموعه‌داده نهایی شامل ۱۹۰۰۰ پرسش چندگزینه‌ای پزشکی به زبان فارسی شد. این مجموعه‌داده، PersianMedQA از پرسش‌های ترجمه‌شده، پرسش‌های جمع آوری شده و پرسش‌های انتخاب شده از ترکیبی از پرسش‌های ترجمه‌شده، پرسش‌های جمع آوری شده و پرسش‌های انتخاب شده از PersianMedQA است و یک منبع ارزشمند برای توسعه مدل‌های زبانی پزشکی فارسی در حوزه پزشکی به شمار می‌آید.

فصل ۵

معرفی مدل زبانی گائوکرنا-V

۱.۵ مقدمه

در این فصل با استفاده از دادگانی که گردآوری کرده ایم یک مدل پایه^۱ را تمرین میدهیم تا مدل جدید مدل گائوکرنا-V را معرفی کنیم که همانطور که در فصل نخست درباره آن صحبت شد دارای توانایی استدلال نیست. در ادامه با استفاده از ترجمه قسمت پژوهشی مجموعه داده U MML به مقایسه مدل جدید خود با مدل پایه و بقیه جایگزین‌ها خواهیم پرداخت.

۲.۵ مدل پایه

به دلیل عدم وجود یک مدل زبانی پژوهشی فارسی متن باز^۲ ما مجبور به انتخاب یک مدل زبانی همه منظوره^۳ به عنوان مدل پایه هستیم.

گزینه‌های متعددی مانند [۵۸] gemma2 ، [۵۷] aya-expanse ، [۵۶] qwen2.5 و PersianMind برای انتخاب در دسترس بودند، که به دلیل زیر مدل aya-expanse-8b انتخاب شده است.
دلیل نخست این است که داده‌های آموزشی سایر مدل‌ها عمدها شامل زبان‌های غیر فارسی هستند، که این امر می‌تواند منجر به ایجاد سوگیری‌هایی در مدل شود که حتی در صورت دستور صریح به استفاده از زبان فارسی،

baseline model^۱
open source^۲
general purpose^۳

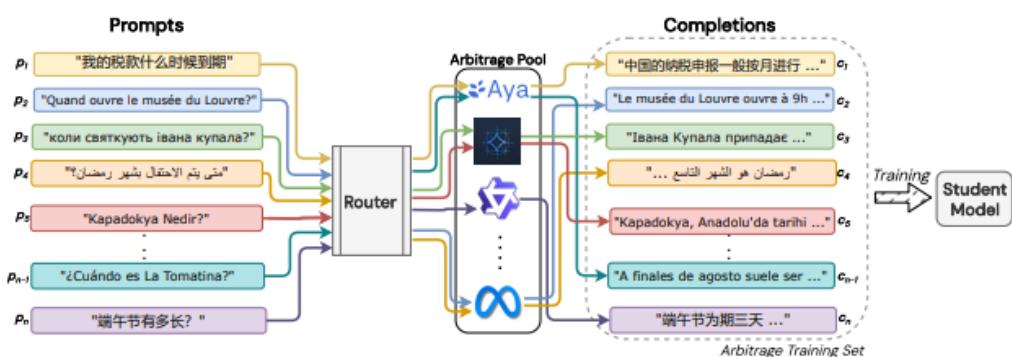
باعث تولید کاراکترهای غیر فارسی می‌شود. در مقابل، aya-expanse درک قوی‌ای از دستور زبان فارسی نشان می‌دهد و متنی غنی و دستوری صحیح به زبان فارسی تولید می‌کند، که آن را به گزینه‌ای بهتر برای پژوهش ما تبدیل می‌کند.

علاوه بر این، اگر ما پارامترهای به روزرسانی شده خود را به جای aya-expanse در مدل دیگری از خانواده aya-vision [۶۰] ادغام کنیم، این امکان را به دست می‌آوریم که تصاویر پزشکی مانند MRI و CT را به عنوان ورودی بپذیریم. این امر باعث افزایش قابلیت کاربرد مدل ما در حوزه پزشکی خواهد شد.

۱.۲.۵ ویژگی‌های مدل aya-expanse

همانطور که پیشتر از آن یاد شد مدل زبانی aya-expanse قابلیت تولید متن‌های فارسی غنی با دستور زبانی صحیح را دارد^۴ ویژگی که بقیه مدل‌های زبانی تنها برای برخی از زبان‌های دارای منابع غنی^۵ مانند انگلیسی و چینی دارا هستند.

این ویژگی زمانی قابل دسترس است که علاوه بر در دسترس بودن منابع غنی برای تمرین منابع به صورت مساوی بین زبان‌هایی که مدل زبانی پشتیبانی می‌کند تقسیم شده باشد؛ aya-expande با تولید دادگان مصنوعی^۶ بر مشکل نبود دادگان فائق آمده است، البته تمرین دادن یک مدل زبانی با دادگان مصنوعی تولید شده توسط یک مدل زبانی دیگر باعث فروپاشی مدل^۷ [۶۱] می‌شود که aya-expande با معرفی مکانیسم آربیتاژ داده [۶۲] که در شکل ۱.۵ می‌بینید از رخداد این اتفاق جلوگیری کرده است.



شکل ۱.۵: مکانیسم آربیتاژ داده

^۴ به جز زبان فارسی این مدل از بیست و دو زبان دیگر نیز پشتیبانی می‌کند.

^۵ resource-rich languages

^۶ synthetic data

^۷ model collapse

۳.۵ تنظیم دقیق روی پیکره پزشکی

برای دستیابی به مدل مدل گائوکرنا-V همان طور که در نمودار کاهش خطای تصویر ۲.۵ میینید ابتدا مدل پایه روی شصت درصد از پیکره پزشکی گردآوری شده تنظیم دقیق^۸ شده است^۹، برای این کار از اندازه دسته^{۱۰} برابر با دو استفاده کردیم تا نیاز به حافظه در طول آموزش کاهش یابد. علاوه بر این، از تجمع گرادیان^{۱۱} با شانزده مرحله استفاده کردیم که به طور مؤثر اندازه کلی دسته را به سی و دو افزایش داد و دینامیک پایدار آموزش را فراهم کرد.

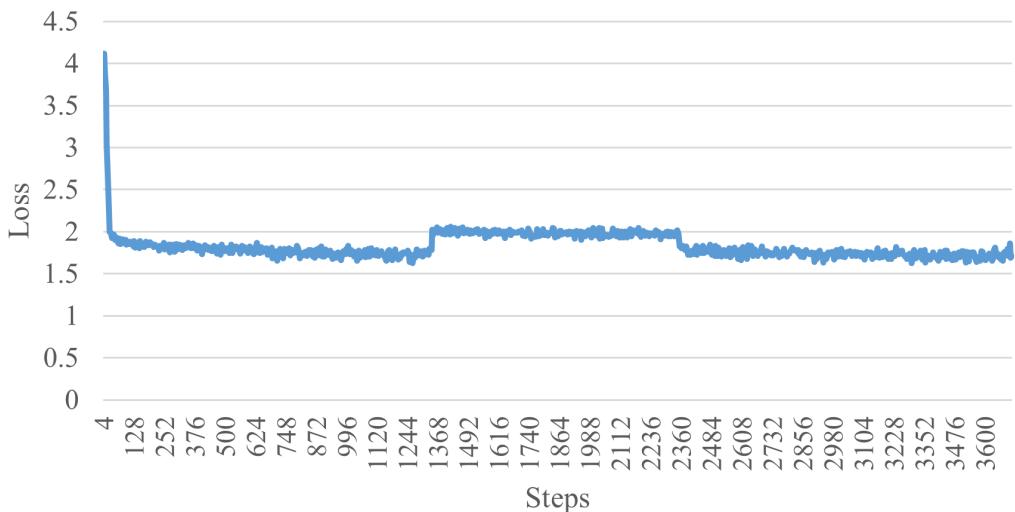
برای کاهش بیشتر مصرف حافظه در فرآیند تنظیم دقیق، از روش LoRA [۶۳]^{۱۲} بهره بردیم تا تعداد پارامترهای قابل آموزش به طور چشمگیری کاهش یابد. برای این کار از رتبه^{۱۳} برابر با هشت، مقدار آلفا برابر با شانزده، نرخ حذف^{۱۴} برابر با پنج درصد و نرخ پوسیدگی وزن^{۱۵} ده درصد استفاده کرده و وزن‌های LoRA را به تمام پارامترهای قابل آموزش در هر لایه ترانسفورمر اختصاص دادیم.

برای بهینه‌سازی بیشتر این فرآیند، از تکنیک‌های کارآمد توکن‌سازی^{۱۶} و تکنیک‌های آموزش مبتنی بر مدیریت حافظه^{۱۷} بهره بردیم. فرآیند توکن‌سازی متن ورودی را به دنباله‌های قابل مدیریت توکن تقسیم کرد، و با کوتاه کردن، پر کردن، و مدیریت توکن‌های اضافی، ساختار ورودی و برچسب‌ها را ثابت نگه داشت تا یکپارچگی مفهومی در طول ثابت زمینه حفظ شود. این آماده‌سازی ساده شده، همراه با تنظیم دقیق مبتنی بر LoRA، و با استفاده از ۲ Flash Attention [۶۴] بهبود یافت.

Flash Attention 2 با کاهش سربار حافظه، امکان مدیریت طول‌های زمینه بلندتر و اندازه‌های دسته بزرگ‌تر را به صورت کارآمد فراهم میکند، که به تنظیم دقیق موثر کمک کرده و تعادل بین کارایی محاسباتی و عملکرد مدل را برقرار میسازد.

^۹ برای دسترسی به کد تنظیم دقیق مدل پایه روی پیکره پزشکی گرد آوری شده میتوانید به آدرس مراجعه کنید <https://github.com/Mehrdadghassabi/Gaokerena-V/blob/main/fine-tuning/Pretraining.ipynb>

fine tune^{۱۸}
batch size^{۱۹}
gradient accumulation^{۲۰}
low rank adaptation^{۲۱}
rank^{۲۲}
drop out^{۲۳}
weight decay^{۲۴}
efficient tokenization^{۲۵}
memory-aware training techniques^{۲۶}



شکل ۲.۵: نمودار کاهش خطای تنظیم دقیق روی پیکره پزشکی

۴.۵ تنظیم دستورالعملی روی مجموعه داده MF3QA

پس از تنظیم دقیق روی پیکره پزشکی، با استفاده از مجموعه داده MF3QA که فصل پیشین معرفی شد انجام دهیم.^{۱۹} همان‌طور که در نمودار کاهش خطای در تصویر ۳.۵ می‌بینید برای این کار به طور مشخص، باز هم از روش LoRA استفاده کرده ایم ولی این بار با رتبه^{۲۰} برابر با دو، آلفا برابر با دو، نرخ حذف برابر با چهل درصد و نرخ پوسیدگی وزن^{۲۱} پنجاه درصد استفاده کرده ایم. فرآیند تنظیم دستورالعمل تنها برای یک دوره^{۲۲} انجام شد تا مدل شیوه پاسخگویی درست را بهتر درک کند.

۵.۵ رد پای کربن مدل گائوکرنا-V

اثر کربنی حاصل از بهینه‌سازی مدل گائوکرنا-V که شامل مراحل تنظیم دقیق و تنظیم دستورالعملی می‌شود، بر اساس مشخصات سخت‌افزاری و مدت زمان اجرا تخمین زده شده است. فرآیند تمرین به مدت نوزده ساعت

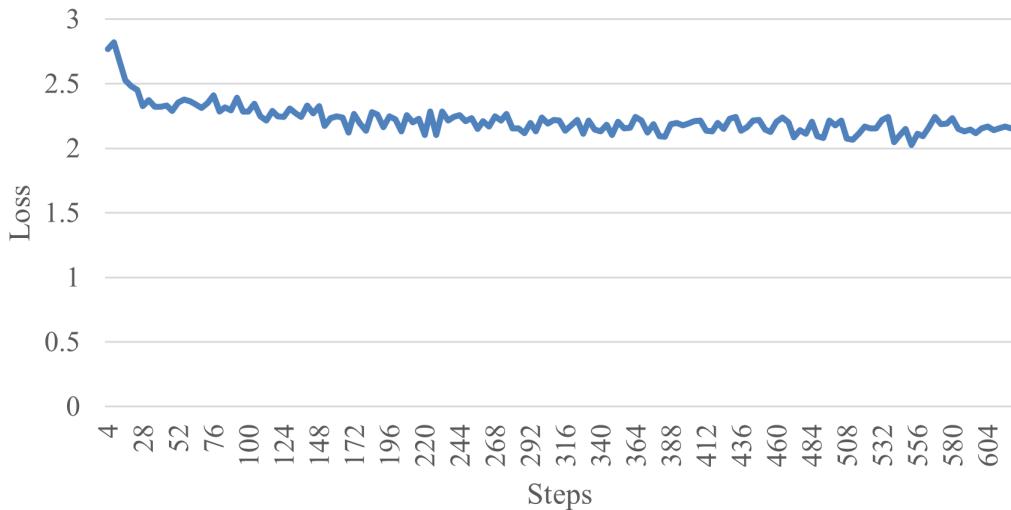
instruction tuning^{۱۸}

برای دسترسی به کد تنظیم دستورالعملی روی مجموعه داده MF3QA می‌توانید به آدرس <https://github.com/Mehrdadghassabi/Gaokerena-V/blob/main/fine-tuning/InstructionTuning.ipynb> مراجعه کنید

rank^{۲۰}

weight decay^{۲۱}

epoch^{۲۲}



شکل ۳.۵: نمودار کاهش خطای تنظیم دقیق روی مجموعه داده MF3QA

بر روی کارت گرافیک NVIDIA A100 PCIe 40GB که در منطقه شرق آسیا^{۲۳} پلتفرم ابری گوگل میزبانی می‌شدند، اجرا شده است.

با توجه به مصرف برق معمولی هر پردازنده گرافیکی که برابر با دویست و پنجاه وات است، کل انرژی مصرف شده در این مدت برابر ۴.۷۵۰ کیلووات ساعت است. با در نظر گرفتن ضریب شدت کربن شبکه شرق آسیا که برابر با ۵۶۰ گرم کربن دی اکسید معادل به ازای هر کیلووات ساعت است، میزان انتشار کربن در طول این فرآیند به ۲۶۶۰ گرم می‌رسد. [۶۵]

۶.۵ نتایج

در نبود مدل‌های پزشکی فارسی برای مقایسه ما مدل گائوکرنا-V را با مدل‌های زبانی فارسی همه منظوره^{۲۴} و جایگزین‌های خط لوله‌ای^{۲۵} مقایسه کرده‌ایم.

همانطور که در شکل ۴.۵ می‌بینید جایگزین‌های خط لوله‌ای شامل یک سری مراحل است ابتدا، یک مترجم پرسش کاربر را از فارسی به انگلیسی تبدیل می‌کند، سپس این پرسش انگلیسی به یک مدل زبانی پزشکی انگلیسی داده می‌شود، و در نهایت، پاسخ تولید شده توسط مدل انگلیسی دوباره از انگلیسی به فارسی ترجمه می‌شود.

asia-east1^{۲۳}
general purpose language models^{۲۴}
pipeline alternatives^{۲۵}

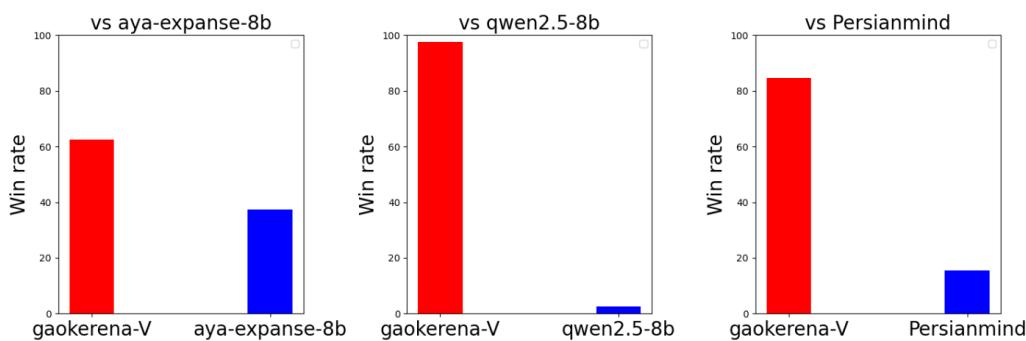


شکل ۴.۵: مکانیسم جایگزین خط لوله ای

۱.۶.۵ مقایسه با مدل های زبانی فارسی همه منظوره

همان طور که در جدول ۱.۵ مشاهده می‌کنید، مدل گانوکرنا-V توانت با موفقیت کنکور علوم پایه پژوهشی شهریور ۱۴۰۲، را پشت سر بگذارد^{۲۶} و به اولین مدل زبان فارسی با کمتر از هشت میلیارد پارامتر تبدیل شد که این آزمون را با موفقیت پشت سر گذاشته است. علاوه بر این، مدل ما در قسمت پژوهشی مجموعه داده^{۲۷} نیز بهبودهایی را نشان داد و نه تنها میانگین نمرات بالاتری کسب کرد، بلکه در اکثر زیرشاخه‌ها عملکرد بر جسته‌ای داشت و توانایی خود را در درک و تولید دانش پژوهشی به زبان فارسی به نمایش گذاشت.

قابل توجه است که دلیل سرعت بالای پاسخگویی PersianMind، همان طور که در جدول ۱.۵ نشان داده شده، این است که این مدل تمایل دارد پاسخ‌های بسیار کوتاه‌تری نسبت به مدل‌های دیگر تولید کند و توکن پایان پاسخ را زودتر ایجاد نماید. علاوه بر ارزیابی پرسش و پاسخ چندگزینه‌ای، ما از GPT-40 [۶۶] به عنوان داور برای پرسش و پاسخ آزاد نیز استفاده کردیم. قسمت تست مجموعه داده MF3QA به مدل زبان رقیب و مدل ما ارائه شد. همان‌طور که در شکل ۵.۵ نشان داده شده است، GPT-40 به طور عمده پاسخ‌های تولیدشده توسط مدل ما را نسبت به سه مدل زبان دیگر ترجیح داده است.



شکل ۵.۵: نرخ پیروزی گانوکرنا-V در رقابت با بقیه مدل های زبانی فارسی همه منظوره

^{۲۶} نمره قبولی در این امتحان سی و شش درصد بدون نمره منفی بوده است.
^{۲۷} MMLU

۲.۶.۵ مقایسه با جایگزین‌های خط لوله‌ای

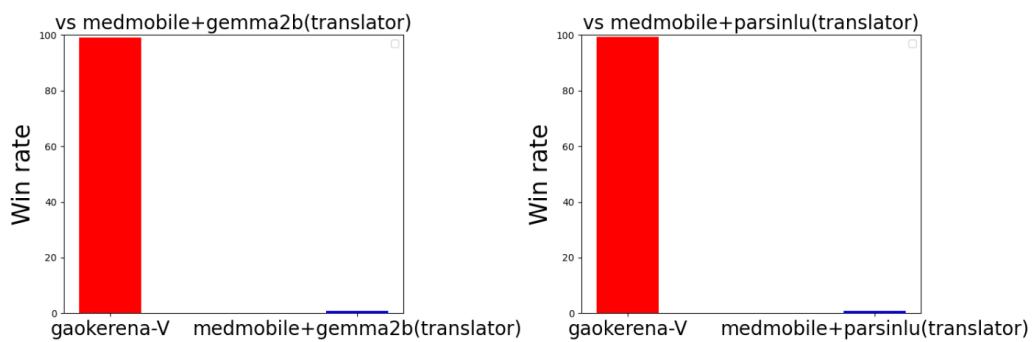
همان‌طور که قبلاً اشاره شد، یکی از گزینه‌های جایگزین برای توسعه یک مدل زبانی پزشکی فارسی، استفاده از جایگزین‌های خط لوله‌ای^{۲۸} است. با این حال، یکی از مشکلات عمدۀ این سیستم‌ها سرعت پایین آن‌ها است. این سیستم‌ها زمان استنتاج بالایی دارند، زیرا خروجی یک مدل باید به مدل دوم منتقل شود و سپس خروجی مدل دوم دوباره توسط مدل اول پردازش شود. این فرآیند تکراری به طور قابل توجهی کارایی سیستم را کاهش می‌دهد.

برای رفع مشکل سرعت پایین سیستم‌های مرحله‌ای، ما تمام پارامترها شامل پارامترهای مربوط به مترجم‌ها و مدل زبانی پزشکی را به طور همزمان بارگذاری کرده‌ایم. آزمایش‌های ما با مدل‌هایی مانند Medmobile همراه با gemma-2b-it به عنوان مترجم، و Medmobile همراه با مدل‌های parsinlu [۶۷] [۶۸] به عنوان مترجم انجام گرفته است؛ همان‌طور که در جدول ۲.۵ مشاهده می‌کنید جایگزین‌های خط لوله‌ای دقیق و سرعت بسیار پایینی از خود نشان دادند.

یکی دیگر از مشکلات مهم جایگزین‌های خط لوله‌ای، عملکرد ضعیف آن‌ها در شناسایی و ترجمه دقیق اصطلاحات پزشکی است. این محدودیت چالشی جدی ایجاد می‌کند، زیرا دقیق در استفاده از اصطلاحات تخصصی برای ارتباط موثر در محیط‌های مراقبت‌های بهداشتی حیاتی است. علت اصلی این ضعف احتمالاً به این دلیل است که مترجم‌های استفاده شده در این سیستم‌ها به طور خاص برای ترجمه پزشکی توسعه نیافرته‌اند. برخلاف مدل‌های ترجمه عمومی، ترجمه پزشکی نیازمند درک دقیق واژگان تخصصی، زمینه و پیچیدگی‌های زبان پزشکی است.

در حال حاضر، هیچ مدلی برای ترجمه پزشکی به زبان فارسی طراحی نشده است، که این امر باعث می‌شود سیستم‌های موجود توانایی کافی برای مدیریت پیچیدگی‌های اصطلاحات پزشکی نداشته باشند. همان‌طور که در شکل ۶ نشان داده شده است، این محدودیت‌ها منجر به نرخ پیروزی بسیار پایین جایگزین‌های خط لوله‌ای در رقابت با مدل گائوکرنا-۷ شده است.

^{۲۸} pipeline alternatives



شکل ۶.۵: نرخ پیروزی گائوکرنا-V در رقابت با جایگزین های خط لوله ای

فصل ۵: معرفی مدل زبانی گائوکرنا-V

توسعه یک مدل زبانی پزشکی مبتنی بر استدلال در زبان فارسی

PersianMind	Qwen2.5	aya-expanse-8b (baseline)	Gaokerena-V (ours)	
25.18	41.48	40.74	48.14	MMLU-anatomy(fa)
34.0	52.0	49.0	53.0	MMLU-medical genetics(fa)
20.23	43.35	44.51	43.93	MMLU-college medicine(fa)
25.28	47.92	52.07	55.47	MMLU-clinical knowledge(fa)
23.89	43.01	45.58	47.05	MMLU-professional medicine(fa)
32.63	44.85	45.14	47.22	MMLU-college biology(fa)
25.89	45.17	46.64	49.31	MMLU(avg)
19.64	33.33	34.52	38.69	IBMSEE Sept 2023
6.8b	7.6b	8b	8b	Number of parameters
2s حدود	15s حدود	8s حدود	10s حدود	inference time

جدول ۱.۵: مقایسه مدل گائوکرنا-V با بقیه مدل های زبانی فارسی همه منظوره

MedMobile parsinlu +	MedMobile + gemma2 -2b-it	Gaokerena (ours)	
25.18	14.07	48.14	MMLU-anatomy(fa)
35.0	20.0	53.0	MMLU-medical-genetics(fa)
27.17	19.08	43.93	MMLU-college-medicine(fa)
31.70	27.54	55.47	MMLU-clinical-knowledge(fa)
33.82	17.27	47.05	MMLU-professional-medicine(fa)
31.25	18.75	47.22	MMLU-college-biology(fa)
30.99	20.11	49.31	MMLU(avg)
32.73	24.40	38.69	IBMSEE Sept2023
3.8b+1.2b+1.2b	3.8b+2b	8b	Number of parameters
حدود 30s	حدود 20s	حدود 10s	inference time

جدول ۲.۵: مقایسه مدل گائوکرنا-V با جایگزین های خط لوله ای

فصل ۶

استدلال در مدل‌های زبانی

ظهور معماری ترنسفورمر^۱ [۵] در سال ۲۰۱۷ نقطه‌ی عطفی بنیادین در تحول مدل‌های زبانی بزرگ بهشمار می‌آید. این معماری با معرفی سازوکار توجه^۲ و جایگزینی آن به جای پردازش متوالی، بسیاری از محدودیت‌های مدل‌های پیشین – از جمله پدیده‌ی محو شدن گرادیان‌ها و ناتوانی در مدل‌سازی وابستگی‌های بلندمدت – را برطرف ساخت. بدین ترتیب، امکان موازی‌سازی کامل در فرایند آموزش فراهم شد و مدل توانست ساختارهای زبانی پیچیده را با دقت و کارایی بسیار بالاتری درک و بازنمایی کند.

در این فصل، تمرکز بر بررسی توانایی استدلال در معماری ترنسفورمر و رویکردهای نوین برای بهبود آن است. هرچند این معماری در وظیفه‌ی پیش‌بینی توکن بعدی عملکردی خیره‌کننده دارد، اما نحوه‌ی استدلال آن عموماً الگومحور است تا منطق محور. به بیان دیگر، تمامی مدل‌های زبانی مدرن که بر پایه‌ی ترنسفورمر ساخته شده‌اند، در هسته‌ی خود از همان ضعف بنیادی رنج می‌برند: ناتوانی در درک و بازتولید فرایندهای استدلالی عمیق و مبتنی بر منطق.

۱.۶ ناتوانی معماری ترنسفورمر در تفکر سیستم دو

در سال ۲۰۱۹، در کنفرانس NeurIPS، یوشوا بنجیو – یکی از چهره‌های پیشگام در حوزه‌ی یادگیری عمیق به یکی از محدودیت‌های بنیادین سیستم‌های یادگیری عمیق امروزی، از جمله مدل‌های زبانی، اشاره کرد: ناتوانی

Transformer^۱
Attention^۲

در انجام وظایفی که نیازمند مهارت‌های استدلالی قوی هستند. [۶۹]^۳ به رغم عملکرد چشمگیر این سیستم‌ها در وظایف شهودی و ادراکی نظریه تشخیص تصویر، ترجمه‌ی ماشینی یا تولید متن آن‌ها در مواجهه با مسائلی که مستلزم استدلال منطقی، درک روابط علی و پردازش شناختی عمیق هستند، عملکرد ضعیفی از خود نشان می‌دهند. در واقع، بسیاری از اشتباهات ساده و احتمانه‌ی مدل‌های زبانی مانند ناتوانی در حل مسائل ریاضی ابتداً یا استنتاج‌های منطقی آشکار را می‌توان با همین محدودیت در تفکر سیستم دو توضیح داد. این مشاهده یادآور تمایز مطرح شده توسط دنیل کانمن در علوم شناختی [۷۰] میان دو نوع تفکر است: تفکر سریع که بنجیو آن را «تفکر سیستم یک» می‌نامد و تفکر آهسته که بنجیو آن را «تفکر سیستم دو» می‌خواند. تفکر سریع، ماهیتی شهودی، واکنشی و خودکار دارد و با نقاط قوت سیستم‌های فعلی هوش مصنوعی هم راست است؛ در مقابل، تفکر آهسته فرایندی آگاهانه، تحلیلی و منطقی است که مستلزم توانایی در برنامه‌ریزی، استدلال و بازنمایی مفاهیم انتزاعی می‌باشد. به بیان دیگر، شبکه‌های عصبی عمیق و بهویژه مدل‌های زبانی مبتنی بر ترنسفورمر عمده‌ای در سطح تفکر سیستم یک، یعنی شناخت شهودی، عمل می‌کنند و فاقد شناخت منطقی یا استدلال علی تفکر سیستم دو هستند. این شکاف میان شهود و منطق، چالشی اساسی در مسیر توسعه‌ی نسل بعدی مدل‌های زبانی و سامانه‌های هوش مصنوعی با قابلیت تفکر مشابه انسان به شمار می‌رود.

۲.۶ طراحی یک معماری دارای قابلیت تفکر سیستم دو

بهترین رویکرد برای غلبه بر ناتوانی ترنسفورمرها در تفکر سیستم دو، توسعه‌ی یک معماری جدید است که بتواند این قابلیت را به‌طور ذاتی در خود داشته باشد. اما طراحی و پیاده‌سازی چنین معماری‌ای کار ساده‌ای نیست! چنین معماری‌ای باید دارای توانایی تعمیم خارج از توزیع^۴ نیز باشد؛ به این معنا که مدل بتواند مفاهیم و روابطی را یاد بگیرد که به صورت صریح و مستقیم در داده‌های آموزشی وجود ندارند. به بیان دیگر، سیستم باید قادر باشد از دانش موجود در داده‌ها به موقعیت‌های جدید و نادیده تعمیم دهد—ویژگی‌ای که برای استدلال واقعی و تفکر سطح بالا ضروری است.

در سال‌های اخیر، پژوهشگران چندین مدل و معماری جایگزین را برای بهبود توانایی استدلال در شبکه‌های عصبی پیشنهاد داده‌اند. برای مثال، یکی از این تلاش‌ها معماری‌ای به نام RIM [۷۱]^۵ بود که توسط یوشوا بنجیو معرفی شد. یا معماری مدل استدلال سلسله‌مراتبی^۶ [۷۲] که با الهام از سلسله مراتبی بودن تفکر انسان پیشنهاد

^۳ سخنرانی بنجیو اهمیت بسزایی داشت و مفهوم «تفکر سیستم دو» را در هوش مصنوعی بر جسته کرد. درس دکتر رهبان در دانشگاه شریف با این عنوان، راهنمایی کلیدی برای پایان‌نامه‌ی حاضر بوده است.

Out-of-Distribution Generalization^۷
Hierarchical Reasoning Model^۸

شده است؛

با این حال، مدل‌های پیشنهاد شده در عمل معماري مقیاس‌پذیری نبوده و نتوانسته اند در مقیاس‌های بزرگ عملکرد قابل قبولی ارائه دهند. به همین دلیل، نمی‌توان آن‌ها را یک پیشنهاد موفق برای حل مشکل استدلال در مدل‌های مبتنی بر ترانسفورمر دانست.

در نتیجه، همچنان یکی از چالش‌های باز در حوزه‌ی یادگیری عمیق و مدل‌های زبانی بزرگ، یافتن معماري‌ای کارا، مقیاس‌پذیر، دارای توان استدلالی و برخوردار از قابلیت تعمیم خارج از توزیع است؛ معماري‌ای که بتواند محدودیت‌های ذاتی ترانسفورمرها را پشت سر بگذارد.

۳.۶ بهبود توانایی دلیل آوری مدل‌های زبانی مبتنی بر ترانسفورمر

تا زمانی که هنوز معماري‌ای با قابلیت استدلال ذاتی^۶ در دسترس نداریم، ناچاریم همچنان با معماري ترانسفورمر کار کنیم و به نوعی بر محدودیت‌ها و ضعف‌های ذاتی آن غلبه نماییم. با وجود تمام پیشرفت‌ها، ترانسفورمرها هنوز قادر سازوکار درونی برای استدلال گام‌به‌گام، تعمیم منطقی و درک علیت هستند. با این حال، از آن‌جا که در حال حاضر موثرترین و مقیاس‌پذیرترین معماري در حوزه‌ی مدل‌های زبانی محسوب می‌شوند، راهبرد عملی این است که به جای کنار گذاشتن آن‌ها، راهکارهایی برای تقویت توانایی استدلالی شان طراحی و به کار گرفته شود. روش‌های گوناگونی برای این هدف پیشنهاد شده است که برای اطلاعات بیشتر میتوانید در مقاله مروی آقای پان [۷۳] در این زمینه مراجعه کنید.

در این بخش، به بررسی مهم‌ترین این رویکردها و روش‌هایی که می‌توانند به بهبود توانایی استدلالی مدل‌های مبتنی بر ترانسفورمر منجر شوند، خواهیم پرداخت.

۱۰.۳.۶ روش‌های مبتنی بر هوش مصنوعی عصبی-نمادین

در حوزه‌ی هوش مصنوعی، دو مکتب فکری اصلی وجود دارد: نمادگرایی^۷ و ارتباطگرایی^۸.

در مکتب نمادگرایی، هدف اصلی این است که منطق، قوانین و روابط صریح میان مفاهیم جمع‌آوری شده و در قالب یک مدل منطقی یا قاعده‌محور نمایش داده شود. این رویکرد بر پایه‌ی استدلال نمادین استوار است؛ یعنی سیستم تلاش می‌کند با استفاده از قواعد صریحی که توسط انسان تعریف شده‌اند، به نتیجه برسد.

Intrinsic Reasoning Capability^۶
Symbolism^۷
Connectionism^۸

در مقابل، مکتب ارتباطگرایی بر پایه‌ی ایده‌ی یادگیری از داده‌ها شکل گرفته است. در این رویکرد، به جای تعریف قوانین منطقی، حجم زیادی از داده به مدل داده می‌شود تا خود سیستم از درون داده‌ها الگوها و روابط را کشف کند. شبکه‌های عصبی مصنوعی و بهویژه مدل‌های عمیق^۹ از نمونه‌های شاخص این مکتب هستند.

با معرفی معماری ترانسفورمر، مکتب ارتباطگرایی به شکل چشمگیری غالب شد و تقریباً تمامی پژوهش‌های پیشرفته‌ی حوزه‌ی هوش مصنوعی حول آن متمرکز گردید. دلیل این غلبه روش‌ن است: جمع‌آوری داده‌های عظیم بسیار ساده‌تر از استخراج و تعریف قوانین منطقی دقیق است، و مدل‌های ارتباطگرای با استفاده از این داده‌های گسترده می‌توانند عملکردی چشمگیر در وظایف گوناگون مانند ترجمه، تولید متن، بینایی ماشین و غیره داشته باشند.

با این حال، همان‌طور که پیش‌تر اشاره شد، ترانسفورمرها ضعف بزرگی دارند: ناتوانی در استدلال! این مدل‌ها اگرچه در شناسایی الگوها و پیش‌بینی توالی‌ها بسیار قدرتمندند، اما فاقد سازوکار درونی برای انجام استدلال منطقی و تفکر نظاممند (تفکر سیستم دو) هستند. از آنجا که تاکنون معماری ای با قابلیت استدلال ذاتی معرفی نشده است، تمامی مدل‌های مبتنی بر ارتباطگرایی نیز همین ضعف بینایی را به ارث برده‌اند. در مقابل، رویکردهای نمادین از این نظر برترند که توانایی استدلال و نتیجه‌گیری منطقی را به صورت ذاتی در خود دارند، اما مشکل بزرگ آن‌ها عدم مقیاس‌پذیری^{۱۰} است؛ یعنی پیاده‌سازی و گسترش آن‌ها برای مسائل واقعی و داده‌های بزرگ بسیار دشوار است.

برای رفع این شکاف، گروهی از پژوهشگران پیشنهاد یک رویکرد میان‌رشته‌ای جدید را مطرح کرده‌اند که ترکیبی از این دو مکتب است: رویکرد عصبی-نمادین^{۱۱}. در این دیدگاه، تلاش می‌شود از قدرت یادگیری و تعمیم مدل‌های عصبی (ارتباطگرای) در کنار توان استدلال و شفافیت سیستم‌های نمادین استفاده شود تا هوشی ترکیبی و متوازن به دست آید؛ هوشی که هم بتواند از داده بیاموزد، و هم به طور منطقی بیندیشد. [۷۴]

با این حال، یوشوا بنجیو معتقد است که این مسیر، راه حل نهایی مناسبی نیست و ممکن است با مقیاس‌پذیری در تضاد باشد. او تأکید می‌کند که به جای ترکیب دو مکتب مجزا، بهترین راه، توسعه‌ی معماری‌هایی است که قابلیت استدلال را به‌طور ذاتی در درون خود داشته باشند؛ یعنی سامانه‌هایی که بتوانند استدلال، تعمیم و یادگیری را در یک چارچوب واحد و منسجم ادغام کنند.

۲.۳.۶ روش‌های مبتنی بر راهنمایی مدل زبانی توسط یک مدل پاداش

یکی از روش‌های بهبود مهارت‌های استدلالی یک مدل زبانی، هدایت فرآیند رمزگشایی آن با استفاده از یک مدل پاداش است. مدل پاداش می‌تواند به دو صورت باشد: مدل پاداش خروجی^{۱۲} و مدل پاداش فرآیند^{۱۳} [۷۵]. مدل پاداش خروجی بر ارزیابی کیفیت پاسخ نهایی تولیدشده توسط مدل زبانی تمرکز دارد. این مدل به خروجی‌ها امتیاز می‌دهد و بر اساس معیارهای مشخص، مانند دقت، انسجام یا ارتباط با پرس‌وجو، آن‌ها را رتبه‌بندی می‌کند. از سوی دیگر، مدل پاداش فرآیند به مراحل میانی استدلال مدل توجه می‌کند. این مدل هر گام از فرآیند تفکر یا حل مسئله را بررسی و ارزیابی می‌کند تا اطمینان حاصل شود که مسیر استدلال منطقی و بهینه است. استفاده از این دو نوع مدل پاداش می‌تواند به بهبود عملکرد مدل زبانی در وظایف پیچیده کمک کند. برای مثال، در حل مسائل ریاضی یا پاسخ به سوالات استدلالی، مدل پاداش فرآیند می‌تواند مدل را به سمت رویکردهای گام‌به‌گام و دقیق‌تر هدایت کند، در حالی که مدل پاداش خروجی کیفیت نهایی پاسخ را تضمین می‌کند. ترکیب این دو مدل می‌تواند به ایجاد پاسخ‌هایی منجر شود که نه تنها درست هستند، بلکه فرآیند رسیدن به آن‌ها نیز شفاف و قابل اعتماد است. یکی از کارهایی که سعی در بهبود توانایی دلیل آوری کرده است مدل Medsss [۳۵] است که در فصل کارهای پیشین در مورد آن توضیحات لازمه داده شده است.

۳.۳.۶ روش‌های مبتنی بر تمرین روی داده‌های دلیل آوری

یکی دیگر از روش‌های بهبود مهارت‌های استدلالی یک مدل زبانی، آموزش آن با داده‌های استدلالی است. با انجام این کار، مدل زبانی می‌تواند یاد بگیرد که چگونه استدلال را تقلید کند یا استدلال‌هایی را که قبل از داده‌های آموزشی خود مشاهده کرده است، بازیابی کند. این رویکرد به مدل کمک می‌کند تا در مواجهه با مسائل پیچیده، پاسخ‌هایی منطقی‌تر و ساختاریافته‌تر ارائه دهد. [۷۶]

در این روش، داده‌های آموزشی شامل نمونه‌هایی از فرآیندهای استدلالی، مانند حل گام‌به‌گام مسائل ریاضی، تحلیل‌های منطقی، یا استدلال‌های مبتنی بر شواهد هستند. با قرار گرفتن در معرض این داده‌ها، مدل زبانی نه تنها یاد می‌گیرد که پاسخ‌های درست تولید کند، بلکه می‌آموزد که چگونه به طور نظاممند به آن پاسخ‌ها برسد. این فرآیند به بهبود توانایی مدل در درک عمیق‌تر پرسش‌ها و ارائه پاسخ‌هایی با کیفیت بالاتر کمک می‌کند. در حقیقت مدل زبانی پایه با تمرین روی داده‌های استدلالی یاد می‌گیرد که فرآیند استدلال را شبیه سازی کند و ما تصور کنیم که در حال دلیل آوری است در حالی که معماری که آن مدل روی آن بنای شده توانایی استدلال را ندارد این پدیده

Output Reward Model^{۱۲}
Process Reward Model^{۱۳}

ای است که به آن توهمندی دلیل آوری^{۱۴} میگویند. [۷۷] [۷۸] یکی از مهم‌ترین کارهایی که این رویکرد را با موفقیت پیاده‌سازی کرده، مدل DeepSeek-R [۳۹] است. این مدل با استفاده از مجموعه داده‌های استدلالی گسترده و متنوع آموزش دیده است و توانسته است عملکرد قابل توجهی در وظایف استدلالی پیچیده از خود نشان دهد. DeepSeek-R با بهره‌گیری از تکنیک‌های پیشرفته در آموزش مدل‌های زبانی، قادر است استدلال‌های منطقی را با دقت بالا تولید کند و حتی در مواردی که نیاز به تفکر چندمرحله‌ای است، عملکردی رقابتی با مدل‌های بزرگ‌تر ارائه دهد. برای اطلاعات بیشتر درباره این مدل و جزئیات رویکرد آن، می‌توانید به بخش کارهای پیشین مراجعه کنید.

فصل ۷

معرفی مدل زبانی گائوکرنا-R

۱.۷ مقدمه

در این بخش، ما به معرفی gaokerena-R می‌پردازیم. در این نسخه، دو روش نوین ارائه می‌شود که با استفاده از مقدار کمی داده می‌توانند توانایی‌های استدلالی مدل پایه را به صورت قابل توجهی بهبود بخشنند. نتایج نشان می‌دهد که این رویکرد جدید عملکردی بهتر از gaokerena-V دارد؛ مدلی که با حجم زیادی از داده‌ها آموزش دیده است. در ادامه، این فرضیه مطرح می‌شود که ارتقای مهارت‌های استدلالی در مدل‌های زبانی کوچک و کم منبع حوزه پژوهشی، تاثیرگذارتر و سودمندتر از صرفا افزایش مقیاس داده‌ها است.

۲.۷ مدل پایه

به همان دلیلی که برای مدل-V، مدل 8b aya-expansione [۵۷] را به عنوان مدل پایه^۱ انتخاب کردیم، برای مدل R gaokerena نیز از همان مدل پایه استفاده می‌کنیم. هدف از این انتخاب، حفظ یک مبنای یکسان برای ارزیابی دقیق و منصفانه میان دو رویکرد است. دلیل دیگر این تصمیم، امکان مقایسه‌ی عادلانه میان آموزش مدل بر حجم عظیمی از داده‌ها و روش پیشنهادی ما برای بهبود مهارت‌های استدلالی است. با این کار، می‌توان به روشنی تاثیر واقعی افزایش توانایی استدلال در برابر صرفا گسترش داده‌های آموزشی را بررسی کرد و نشان داد که چگونه تقویت استدلال می‌تواند حتی در شرایط داده‌های محدود، موجب ارتقای عملکرد مدل شود.

¹ baseline model

به دلیل محدودیت‌های سخت افزاری و سرعت بالای چارچوب اول، مدل پایه aya-expanse-8b را با چارچوب اول برای 95 درصد داده‌ها و چارچوب دوم برای 5 درصد باقی مانده آموزش دادیم. این رویکرد، آموزش کارآمد با بهینه‌سازی ترجیح مستقیم را ممکن ساخت و شامل 11000 جفت پاسخ ترجیحی-ردشده با حدود 2 میلیون توکن ترجیحی و 2.5 میلیون توکن ردشده بود. این ترکیب، زمان آموزش را بهینه کرد و با بهره‌گیری از بازخورد دقیق چارچوب دوم، مهارت‌های استدلالی مدل را در مسائل پزشکی تقویت نمود.

۳.۷ متد های معرفی شده

همان‌طور که پیش‌تر اشاره شد، برای تقویت مهارت‌های استدلالی در حوزه پزشکی، دو روش نوین معرفی شده‌اند. روش نخست سریع‌تر از روش دوم عمل می‌کند، اما روش دوم داده‌های باکیفیت‌تری برای آموزش تولید می‌نماید. در ادامه، به توضیح جزئیات این دو روش می‌پردازیم.

۱.۳.۷ روش نخست

روش نخست، که در شکل ۱.۷ و الگوریتم ۲.۷ نمایش داده شده است، از یک مدل معلم (در این پژوهش از مدل زبانی DeepSeek-R استفاده شده است) بهره می‌گیرد. این مدل معلم، خود یک مدل زبانی با قابلیت استدلال پیشرفته است که وظیفه دارد خطاهای استدلالی مدل دانش‌آموز را در پاسخ به پرسش‌های چندگزینه‌ای حوزه پزشکی شناسایی و اصلاح کند. در این فرآیند، اگر مدل دانش‌آموز گزینه‌ی صحیح را انتخاب کند، بدون نیاز به مداخله‌ی بیشتر، به سؤال بعدی در مجموعه‌داده منتقل می‌شویم. اما در صورتی که مدل دانش‌آموز پاسخ نادرست دهد، مدل معلم با در نظر گرفتن پاسخ درست، یک توضیح گام‌به‌گام و دقیق در قالب زنجیره استدلال^۲ تولید می‌کند که منطق، شواهد و روند تفکر پشت پاسخ صحیح را آشکار می‌سازد. این توضیحات نه تنها اشتباه مدل دانش‌آموز را مشخص می‌کنند، بلکه نحوه‌ی استدلال صحیح در مسائل پزشکی را نیز آموزش می‌دهند. در این سازوکار، خروجی مدل معلم به عنوان پاسخ برگزیده^۳ تلقی می‌شود و پاسخ اولیه‌ی مدل دانش‌آموز به عنوان پاسخ مردود^۴ برای فرآیند بهینه‌سازی مستقیم ترجیحات^۵ [۱۹] مورد استفاده قرار می‌گیرد. این چارچوب آموزشی به مدل دانش‌آموز امکان می‌دهد تا از طریق بازخورد مستقیم مدل معلم، به شکلی هدفمند مهارت‌های استدلالی

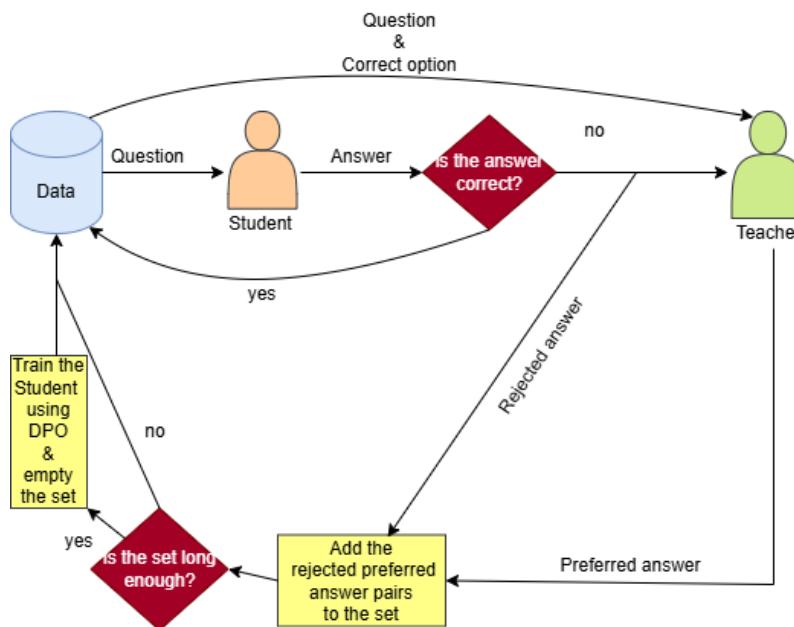
chain-of-thought^۲

preferred answer^۳

rejected answer^۴

direct preference optimization^۵

خود را تقویت کرده و درک عمیق‌تری از منطق پزشکی در پرسش‌های چندگزینه‌ای به دست آورد.



شکل ۱.۷: بلوک دیاگرام شیوه نخست

۲.۳.۷ روش دوم

روش دوم، که در شکل ۲.۷ و الگوریتم ۲.۷ نشان داده شده است، از مدل معلم استفاده می‌کند تا بازخورد دقیق‌تری ارائه دهد. این روش با شناسایی خطاهای خاص در پاسخ مدل دانش‌آموز به یک پرسش چندگزینه‌ای عمل می‌کند. پس از دریافت این بازخورد، مدل دانش‌آموز تشویق می‌شود تا پاسخ خود را دقیقاً از نقطه‌ای که خطأ رخ داده است، اصلاح کند. از طریق این فرآیند تکراری، مدل دانش‌آموز به تدریج مسیر استدلالی^۶ [۷۹] خود را بهبود می‌بخشد و به پاسخ صحیح همگرا می‌شود. این بازخورد سازنده به مدل دانش‌آموز امکان می‌دهد تا استقلال استدلالی خود را تقویت کند، زیرا یاد می‌گیرد خطاهای خود را به جای تقلید صرف از پاسخ‌های مدل معلم، اصلاح نماید. در این چارچوب، پاسخ نهایی تولیدشده توسط مدل دانش‌آموز به عنوان پاسخ ترجیحی در نظر گرفته می‌شود، در حالی که تلاش اولیه آن به عنوان پاسخ رشدشده برای آموزش بهینه سازی مستقیم ترجیحات استفاده می‌شود. اگرچه این چارچوب دوم به دلیل فرآیند بازخورد تکراری زمان‌برتر است، اما آموزش بهینه سازی مستقیم ترجیحات را ساده‌تر می‌کند، زیرا هر دو پاسخ رشدشده و ترجیحی به طور کامل توسط خود مدل دانش‌آموز کمک می‌کند

^۶ reasoning trajectories^۶

الگوریتم ۱.۷ شیوه دوم پیشنهاد شده برای بهبود دلیل آوری

- ورودی: پرسش و پاسخ های چند گزینه پزشکی
- ۱: برای به ازای تمامی پرسش ها انجام بده
 - ۲: یکی از پرسش های پزشکی را از مدل دانشجو بپرس
 - ۳: اگر اگر پاسخ دانشجو درست بود آنگاه
 - ۴: این پرسش را کنار بگذار و به سراغ پرسش بعدی برو
 - ۵: پایان شرط اگر
 - ۶: پاسخ نادرست دانشجو را ذخیره کن
 - ۷: پاسخ نادرست دانشجو را به همراه کلید پاسخ معلم بده و از او بخواه که استدلال نادرست دانشجو را اصلاح کند
 - ۸: پاسخ معلم را به عنوان پاسخ صحیح ذخیره کن
 - ۹: پاسخ درست و نادرست را به عنوان جفت ذخیره کن
 - ۱۰: اگر در صورت جمع شدن تعداد کافی جفت پاسخ درست و نادرست آنگاه
 - ۱۱: مدل دانشجو را با استفاده از بهینه سازی مستقیم ترجیحات تمرين بده
 - ۱۲: پایان شرط اگر
 - ۱۳: پایان حلقة برای

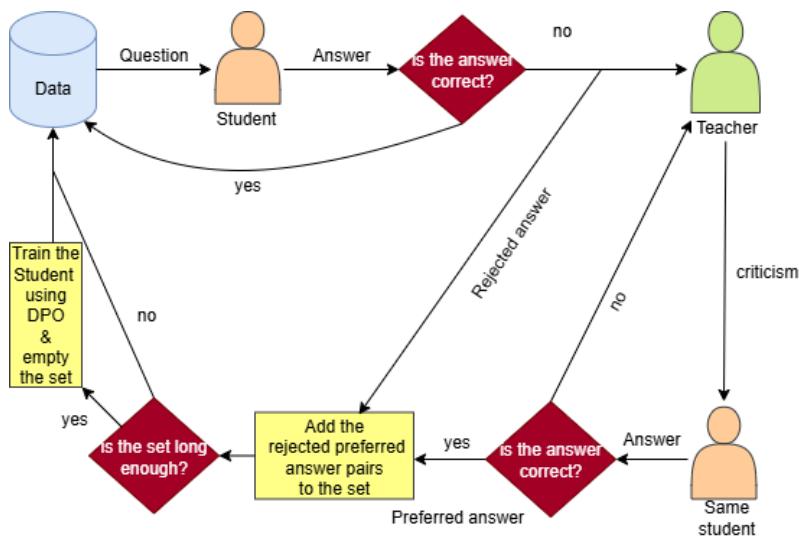
تا مهارت‌های استدلالی خود را به صورت مستقل توسعه دهد، بلکه با تولید داده‌های آموزشی با کیفیت‌تر، کارایی و فرآیند یادگیری را نیز افزایش می‌دهد.

۴.۷ ردپای کربن

ردپای کربنی فرآیند تمرين مدل gaokerena-R بر اساس پیکربندی سخت‌افزاری و مدت‌زمان کل اجرا برآورد شده است. در این فرآیند، با بهره‌گیری از مجموعه‌داده‌ی ترجمه‌شده‌ی ماشینی MedMCQA [۵۱] و به کارگیری دوروش پیشنهادی همراه با بهینه‌سازی مستقیم ترجیحات مدل مورد نظر به مدت تقریبی یک ساعت بر روی یک کارت گرافیک NVIDIA H100 PCIe با حافظه‌ی 80 گیگابایت آموزش داده شد. در طول این فرآیند، حدود 43 گیگابایت از حافظه‌ی گرافیکی^۷ مورد استفاده قرار گرفت. نمودار تغییرات تابع خطأ در طول آموزش در شکل ۳.۷ نمایش داده شده است.

با در نظر گرفتن میانگین توان مصرفی حدود 350 وات برای هر GPU، میزان کل انرژی مصرفی در این فرآیند

VRAM^V



شکل ۲.۷: بلوک دیاگرام شیوه دوم

تقریباً ۰.۳۵ کیلووات ساعت برآورد می‌شود. با توجه به شدت متوسط کربن شبکه‌ی برق کانادا، جایی که سرور ما در آن قرار داشت (هشتاد و شش گرم دی اکسید کربن معادل برای هر کیلووات ساعت)، میزان انتشار کربن حاصل از این فرآیند در حدود سی گرم دی اکسید کربن تخمین زده می‌شود.

در مقایسه با مدل قبلی ما، gaokerena-V، که انتشار تقریبی ۲.۶۶ کیلوگرم دی اکسید کربن داشت، این مقدار کاهش چشمگیری در تاثیرات زیست محیطی را نشان می‌دهد. این نتیجه بیانگر آن است که روش جدید ما نه تنها از نظر کارایی و بهینه‌سازی استدلال عملکرد بهتری دارد، بلکه از منظر پایداری زیست محیطی نیز بسیار کارآمدتر و سازگارتر با محیط زیست عمل می‌کند. [۶۵]

۵.۷ نتایج

در این بخش، مدل تازه توسعه یافته gaokerena-R را با مدل پیشین، gaokerena-V، مقایسه می‌کنیم. در حالی که gaokerena-V روی مجموعه داده‌های پزشکی بزرگی آموزش دیده و دانش پزشکی قوی و قابلیت‌های بازیابی خوبی را نشان می‌دهد، gaokerena-R به طور خاص برای تقویت استدلال پزشکی^۸ طراحی شده است. به دلیل خطمنشی آموزشی مرکزی بر استدلال، gaokerena-R روی مجموعه داده‌ای به مرتب کوچک‌تر آموزش دیده که منجر به پوشش کمی محدودتر دانش پزشکی عمومی شده است. با این حال، مدل gaokerena-R توانسته شایستگی استدلالی عمیق‌تری نشان داده و در وظایفی که نیاز به استنتاج چندمرحله‌ای و انسجام منطقی دارند،

^۸ medical reasoning^۸

الگوریتم ۲.۷ شیوه دوم پیشنهاد شده برای بهبود دلیل آوری

ورودی: پرسش و پاسخ های چند گزینه پزشکی

۱: برای به ازای تمامی پرسش ها انجام بده

۲: یکی از پرسش های پزشکی را از مدل دانشجو بپرس

۳: اگر اگر پاسخ دانشجو درست بود آنگاه

۴: این پرسش را کنار بگذار و به سراغ پرسش بعدی برو

۵: پایان شرط اگر

۶: پاسخ نادرست دانشجورا ذخیره کن

۷: تا زمانی که تا زمانی که دانشجو پاسخ نادرست را نیافته است انجام بده

۸: پاسخ نادرست دانشجو را به معلم بده و از او بخواه که استدلال نادرست را گزارش کند

۹: گزارش معلم را به دانشجو بده و از او بخواه به پرسش دوباره پاسخ دهد

۱۰: پایان حلقه تا زمانی که

۱۱: پاسخ درست دانشجورا ذخیره کن

۱۲: پاسخ درست و نادرست را به عنوان جفت ذخیره کن

۱۳: اگر در صورت جمع شدن تعداد کافی جفت پاسخ درست و نادرست آنگاه

۱۴: مدل دانشجورا با استفاده از بهینه سازی مستقیم ترجیحات تمرین بده

۱۵: پایان شرط اگر

۱۶: پایان حلقه برای

عملکرد بهتری داشته باشد. در ارزیابی نهایی، عملکرد gaokerena-V با استفاده از درخواست های^۹ مستقیم با

عملکرد gaokerena-R هنگام استفاده از درخواست های زنجیره ای فکری^{۱۰} مقایسه می شود.

نتایج نشان می دهد که gaokerena-R ، با وجود مقیاس کوچک تر و داده های آموختی محدود، از طریق

هدایت استدلال ساختار یافته، عملکرد برتری را نشان داده و این موضوع اثربخشی بهینه سازی مت مرکز بر استدلال

را در مقایسه با افزایش صرف مقیاس داده ها نشان می دهد.

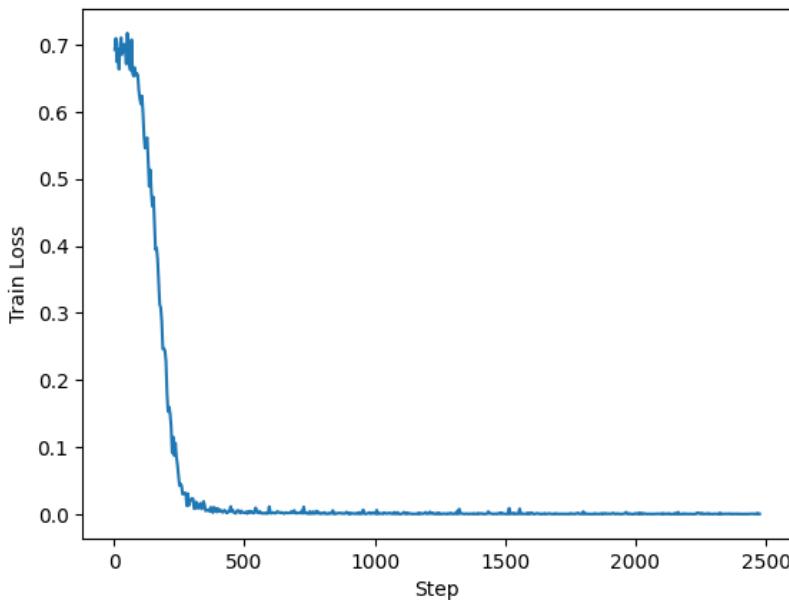
۱.۵.۷ سنجش استدلال پزشکی

برای ارزیابی قابلیت های استدلال پزشکی مدل ها، از آن ها خواستیم تا مسیر های استدلال زنجیره ای فکری

را با تنظیم دمای 1.0 تولید کنند. از آنجا که همه مدل ها ممکن است پاسخ های متفاوتی برای یک پرسش یکسان

ارائه دهند، عملکرد آن ها را با استفاده از دو معیار دقت و pass@k [۸۰] روی مجموعه داده های ترجمه فارسی

^۹ prompt
^{۱۰} chain of thought



شکل ۳.۷: نمودار کاهش خطابهینه سازی مستقیم ترجیحات

MMLU [۱۱] و کنکور علوم پایه پزشکی شهریور ۱۴۰۱ ارزیابی شده است.

۱.۱.۵.۷ دقت

معیار اول، دقت برای هر پرسش در مجموعه داده‌ها است. برای هر پرسش، پنج نمونه مستقل تولید شد و مکانیزم رای‌گیری اکثریت اعمال گردید: اگر سه یا بیشتر از پنج پاسخ تولیدشده گزینه یکسانی را انتخاب کردند، آن گزینه به عنوان پیش‌بینی نهایی انتخاب شد؛ در غیر این صورت، پرسش بدون پاسخ باقی ماند تا عدم قطعیت مدل را نشان دهد. این چارچوب، برآورد قوی‌ای از سازگاری استدلالی در مسیرهای مختلف ارائه می‌دهد. نتایج در جدول ۱.۷ (بدون نمره منفی) و جدول ۲.۷ (با نمره منفی) ارائه شده‌اند، جایی که در تنظیم نمره منفی، هر سؤال پاسخ‌داده شده به صورت نادرست، امتیاز -0.33 دریافت می‌کند.

۲.۱.۵.۷ Pass@k

معیار Pass@k که ابتدا توسط براون و همکاران معرفی شد، یک معیار قوی برای ارزیابی مدل‌های زبانی ارائه می‌دهد که برای یک ورودی یکسان، در نمونه‌گیری‌های متعدد، خروجی‌های متفاوتی تولید می‌کنند. این معیار، احتمال اینکه مدل حداقل یک پاسخ درست را در k تلاش مستقل تولید کند، محاسبه می‌کند و بدین ترتیب ارزیابی جامع‌تری از قابلیت اطمینان استدلال و تنوع نمونه‌ها ارائه می‌دهد. تعریف رسمی این معیار در فرمول

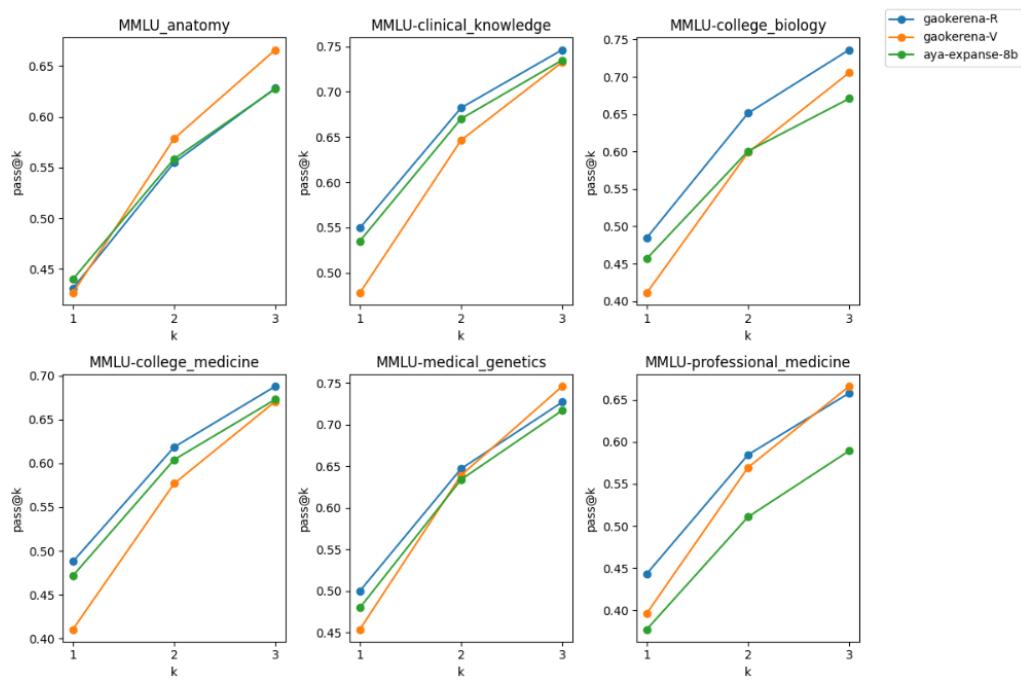
۱.۷ آمده است، که در آن N نشان‌دهنده تعداد کل نمونه‌های تولید شده و C_i نشان‌دهنده تعداد نمونه‌های درست برای مسئله i است.

$$\text{pass@k} = \frac{1}{\# \text{ of problems}} \sum_{i=1}^{\# \text{ of problems}} \left(1 - \frac{\binom{N-C_i}{k}}{\binom{N}{k}} \right) \quad (1.7)$$

با پیروی از همان تنظیمات آزمایشی که پیش‌تر توضیح داده شد، امتیازهای $\text{Pass}@k$ را با $k = 1, 2, 3$ روی دو مجموعه داده ترجمه قسمت پزشکی MMLU و کنکور علوم پایه پزشکی شهریور ۱۴۰۱ محاسبه کردیم. ارزیابی شامل سه مدل گائوکرنا-R، گائوکرنا-8b و aya-expansse-8b بود. نتایج مربوط به مجموعه داده ترجمه قسمت پزشکی MMLU در شکل ۴.۷ و نتایج کنکور علوم پایه پزشکی شهریور ۱۴۰۱ در شکل ۵.۷ نشان داده شده‌اند. همان‌طور که در شکل‌های ۴.۷ و ۵.۷ مشاهده می‌شود، مدل گائوکرنا-R به طور مداوم در تقریباً تمام مقادیر k و دسته‌های ارزیابی، عملکرد برتری از خود نشان داد. این بهبود حاکی از آن است که گائوکرنا-R از قابلیت‌های استدلالی پایدارتر و منسجم‌تری برخوردار است و حتی با تعداد نمونه‌های محدود، پاسخ‌های درست را با اطمینان بیشتری تولید می‌کند. در مقابل، مدل گائوکرنا-8b در مقادیر کوچک k عملکرد نسبتاً ضعیفی داشت، اما با افزایش k نتایج آن بهبود یافت. این الگو نشان می‌دهد که گائوکرنا-8b هنگام تولید مسیرهای استدلالی زنجیره‌ای افکار به طور قابل توجهی نامطمئن‌تر عمل می‌کند و اغلب در نمونه‌های مختلف، پاسخ‌های متعدد یا ناسازگاری تولید می‌نماید. بنابراین نتایج، برتری چشمگیر در ثبات و قابلیت اطمینان استدلال را که از طریق رویکرد آموزشی هدفمند و متمرکز بر استدلال در مدل گائوکرنا-R به دست آمده است، به خوبی برجسته می‌کنند.

۲.۵.۷ سنجش دانش پزشکی

همان‌طور که پیش‌تر اشاره شد، مدل گائوکرنا-8b دارای پایه دانش پزشکی گسترده‌تری نسبت به مدل جدیدتر گائوکرنا-R است؛ عمدتاً به این دلیل که بر روی مجموعه داده‌های پزشکی به مراتب بزرگ‌تر و متنوع‌تری آموزش دیده است. این تفاوت بهویژه زمانی آشکار می‌شود که هر دو مدل با درخواست‌های مستقیم ارزیابی شوند، نه زمانی که از آن‌ها درخواست زنجیره افکار می‌شود؛ زیرا در حالت دوم، توانایی استدلال مدل نیز بر عملکرد نهایی تاثیر قابل توجهی می‌گذارد. همان‌طور که در جدول ۳.۷ نشان داده شده، در شرایط درخواست مستقیم، گائوکرنا-8b عملکرد برتری از خود نشان می‌دهد که این امر بیانگر توانایی قوی‌تر آن در به‌خاطرسپردن و بازیابی حقایق پزشکی ناشی از پیش‌آموزش گسترده است. نکته جالب توجه اینکه عملکرد گائوکرنا-R و مدل پایه آن aya-expansse-8b در شرایط درخواست مستقیم تقریباً برابر است. این مشاهدت حاکی از آن است که برتری



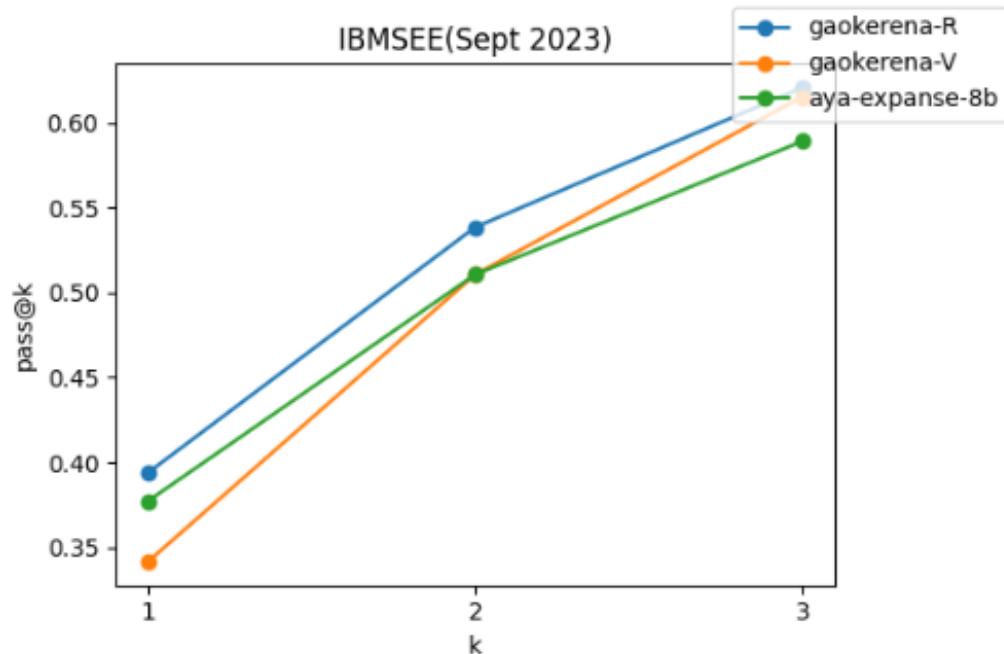
شکل ۴.۷: نتایج Pass@k روی مجموعه‌داده ترجمه قسمت پزشکی MMLU

گائوکرنا-R در سناریوهای زنجیره افکار، عمدتاً ناشی از مهارت‌های استدلالی پیشرفته‌تر آن نسبت به مدل پایه است، نه افزایش قابل توجه دانش پزشکی.

۳.۵.۷ سنجش پایانی

همان‌طور که پیش‌تر بحث شد، مدل گائوکرنا-V در حالت درخواست مستقیم عملکرد بهتری از خود نشان می‌دهد، در حالی که گائوکرنا-R زمانی برتر است که از آن درخواست زنجیره افکار شود؛ یعنی از مدل خواسته شود پیش از ارائه پاسخ نهایی، گام‌به‌گام استدلال کند. یکی از مزایای کلیدی پرامپت کردن با زنجیره افکار - بهویژه هنگامی که با نمونه‌برداری چندگانه^{۱۱} و رای‌گیری اکثریت^{۱۲} ترکیب شود - این است که به‌طور طبیعی میزان اطمینان مدل را نشان می‌دهد. اگر پاسخ‌های تولیدشده در نمونه‌های مختلف به یک گزینه همگرا شوند، می‌توان مدل را مطمئن دانست؛ در مقابل، پراکندگی زیاد پاسخ‌ها نشان‌دهنده عدم اطمینان مدل است. این روش جایگزین عملی و مؤثری برای پرس‌وجو مستقیم از «اطمینان خودآزارزیابی شده مدل» به شمار می‌رود؛ قابلیتی که مدل‌های کوچک‌تر معمولاً به دلیل محدودیت در خودآگاهی و توانایی استدلال درون‌نگرانه از آن بی‌بهره‌اند. [۸۱]

multiple sampling^{۱۱}
majority voting^{۱۲}



شکل ۷.۵: نتایج Pass@k روی مجموعه داده کنکور علوم پایه پزشکی شهریور ۱۴۰۱

در مواردی که گائوکرنا-R عدم اطمینان نشان می‌دهد (یعنی چندین پاسخ متمایز تولید می‌کند)، از مدل پایه یعنی aya-expanse-8b به عنوان یک تاییدکننده کمکی^{۱۳} استفاده می‌کنیم. پاسخ‌های تولیدشده توسط گائوکرنا-R به aya-expanse-8b ارائه می‌شود و این مدل مأمور می‌شود گزینه‌ای را انتخاب کند که کمترین اطلاعات نادرست یا ناسازگار را داشته باشد.

در نتیجه، دو پیکربندی زیر را با هم مقایسه کردیم که در جدول ۴.۷ خلاصه شده‌اند:

- گائوکرنا-V فقط با درخواست مستقیم و بدون هیچ‌گونه راهنمایی استدلالی ارزیابی می‌شود.
- به گائوکرنا-R درخواست زنجیره افکار داده می‌شود و در صورت بروز عدم اطمینان، پاسخ‌های تولیدشده آن توسط aya-expanse-8b بررسی و معتبرترین پاسخ انتخاب می‌شود.

aya-expanse-8b (baseline)	gao kerena-V	gao kerena-R	
40.74	39.25	42.22	MMLU-anatomy(fa)
45.0	41.0	50.0	MMLU-medical-genetics(fa)
48.55	37.57	47.97	MMLU-college-medicine(fa)
54.71	46.79	55.84	MMLU-clinical-knowledge(fa)
43.75	37.13	44.85	MMLU-professional-medicine(fa)
43.75	36.80	48.61	MMLU-college-biology(fa)
47.10	40.40	48.76	MMLU(avg)
35.71	29.76	38.69	IBMSEE Sept2023
8b	8b	8b	of Number parameters
$\approx 5 \times 35s$	$\approx 5 \times 35s$	$\approx 5 \times 35s$	time inference

جدول ۱.۷: دقت با درخواست زنجیره افکار بدون نمره منفی

aya-expansive-8b (baseline)	gao kerena-V	gao kerena-R	
24.93	27.65	29.13	MMLU-anatomy(fa)
33.0	32.0	40.0	MMLU-medical-genetics(fa)
34.48	25.24	34.68	MMLU-college-medicine(fa)
42.51	35.59	44.65	MMLU-clinical-knowledge(fa)
30.39	25.0	30.39	MMLU-professional-medicine(fa)
30.09	25.0	36.80	MMLU-college-biology(fa)
33.55	28.57	36.14	MMLU(avg)
19.84	15.87	24.60	IBMSEE Sept2023
8b	8b	8b	of Number parameters
$\approx 5 \times 35s$	$\approx 5 \times 35s$	$\approx 5 \times 35s$	time inference

جدول ۷: دقت با درخواست زنجیره افکار با نمره منفی

aya-expanse-8b (baseline)	gao kerena-V	gao kerena-R	
40.74	48.14	41.48	MMLU-anatomy(fa)
49.0	53.0	49.0	MMLU-medical-genetics(fa)
44.51	43.93	46.24	MMLU-college-medicine(fa)
52.07	55.47	52.45	MMLU-clinical-knowledge(fa)
45.58	47.05	41.91	MMLU-professional-medicine(fa)
45.14	47.22	44.44	MMLU-college-biology(fa)
46.64	49.31	46.28	MMLU(avg)
34.52	38.69	35.11	IBMSEE Sept2023
8b	8b	8b	of Number parameters
≈ ۱۰s	≈ ۱۰s	≈ ۱۰s	time inference

جدول ۳.۷: دقت با درخواست مستقیم

gaokerena-V	gaokerena-R + aya-expans-8b (verifier)	
48.14	47.40	MMLU- anatomy(fa)
53.0	56.0	MMLU- medical-genetics(fa)
43.93	50.28	MMLU- college-medicine(fa)
55.47	58.86	MMLU- clinical-knowledge(fa)
47.05	48.89	MMLU- professional- medicine(fa)
47.22	54.86	MMLU- college-biology(fa)
49.31	52.98	MMLU(avg)
38.69	46.42	IBMSEE Sept2023
Straight	COT for the main model Straight for the verifier	prompt
$\approx 10s$	$\approx 5 \times 35 + 8s$	time inference

جدول ۴.۷: ارزیابی دو پیکربندی

فصل ۸

نتیجہ گیری

كتاب نامه

and Privacy Data Health on Intelligence Artificial of "Implications Ahmad. Momani. [١]
. (٢٠٢٥) . ١٦٣٩. ٢٥٠ ١arXiv: preprint arXiv Confidentiality."

Brown, Peter F., et al. "A statistical approach to machine translation." (1990): 79-85. [٢]

Williams, Ronald J., and David Zipser. "A learning algorithm for continually running [٣] fully recurrent neural networks." Neural computation 1.2 (1989): 270-280.

Hochreiter, Sepp, and Jürgen Schmidhuber. "Long short-term memory." Neural com- [٤] putation 9.8 (1997): 1735-1780.

Vaswani, Ashish, et al. "Attention is all you need." Advances in neural information [٥] processing systems 30 (2017).

Koroteev, Mikhail V. "BERT: a review of applications in natural language processing [٦] and understanding." arXiv preprint arXiv:2103.11943 (2021).

Devlin, Jacob, et al. "Bert: Pre-training of deep bidirectional transformers for lan- [٧] guage understanding." Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technolo- gies, volume 1 (long and short papers). 2019.

Yenduri, Gokul, et al. "Generative pre-trained transformer: A comprehensive review [٨] on enabling technologies, potential applications, emerging challenges, and future di- rections." arXiv preprint arXiv:2305.10435 (2023).

Raffel, Colin, et al. "Exploring the limits of transfer learning with a unified text-to-text transformer." *Journal of machine learning research* 21.140 (2020): 1-67.

Luo, Man, et al. "Choose your QA model wisely: A systematic study of generative and extractive readers for question answering." arXiv preprint arXiv:2203.07522 (2022).

Hendrycks, Dan, et al. "Measuring massive multitask language understanding." arXiv preprint arXiv:2009.03300 (2020).

Liu, Yang, et al. "G-eval: NLG evaluation using gpt-4 with better human alignment." arXiv preprint arXiv:2303.16634 (2023).

Manes, Itay, et al. "K-qa: A real-world medical q&a benchmark." arXiv preprint arXiv:2401.14493 (2024).

Zhang, Tianyi, et al. "Bertscore: Evaluating text generation with bert." arXiv preprint arXiv:1904.09675 .(2019)

Christiano, Paul F., et al. "Deep reinforcement learning from human preferences." Advances in neural information processing systems 30 .(2017)

Lee, Harrison, et al. "Rlaif vs. rlhf: Scaling reinforcement learning from human feedback with ai feedback." arXiv preprint arXiv:2309.00267 .(2023)

Schulman, John, et al. "Proximal policy optimization algorithms." arXiv preprint arXiv:1707.06347 .(2017)

Shao, Zhihong, et al. "Deepseekmath: Pushing the limits of mathematical reasoning in open language models." arXiv preprint arXiv:2402.03300 .(2024)

Rafailov, Rafael, et al. "Direct preference optimization: Your language model is secretly a reward model." Advances in neural information processing systems 36 (2023): .53728-53741

Ji, Kaixuan, et al. "Enhancing multi-step reasoning abilities of language models [۲۰] through direct q-function optimization." arXiv preprint arXiv:2410.09302 .(2024)

Haarnoja, Tuomas, et al. "Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor." International conference on machine learning. Pmlr, .2018

Singhal, Karan, et al. "Toward expert-level medical question answering with large [۲۲] language models." Nature Medicine (2025): 1-8.

Singhal, Karan, et al. "Large language models encode clinical knowledge." Nature [۲۳] 620.7972 (2023): 172-180.

Saab, Khaled, et al. "Capabilities of gemini models in medicine." arXiv preprint [۲۴] arXiv:2404.18416 (2024).

Li, Yunxiang, et al. "Chatdoctor: A medical chat model fine-tuned on a large language model meta-ai (llama) using medical domain knowledge." Cureus 15.6 (2023).

Touvron, Hugo, et al. "Llama: Open and efficient foundation language models." [۲۶] arXiv preprint arXiv:2302.13971 (2023).

Kim, Hyunjae, et al. "Small language models learn enhanced reasoning skills from [۲۷] medical textbooks." arXiv preprint arXiv:2404.00376 (2024).

Vishwanath, Krithik, et al. "MedMobile: A mobile-sized language model with [۲۸] expert-level clinical capabilities." arXiv preprint arXiv:2410.09019 (2024).

Abdin, Marah, et al. "Phi-3 technical report: A highly capable language model locally [۲۹] on your phone." arXiv preprint arXiv:2404.14219 (2024).

Taghizadeh, Nasrin, et al. "SINA-BERT: a pre-trained language model for analysis [۳۰] of medical texts in Persian." arXiv preprint arXiv:2104.07613 (2021).

Koroteev, Mikhail V. "BERT: a review of applications in natural language processing [۳۱] and understanding." arXiv preprint arXiv:2103.11943 (2021).

Veisi, Hadi, and Hamed Fakour Shandi. "A Persian medical question answering sys- [۳۲] tem." International Journal on Artificial Intelligence Tools 29.06 (2020): 2050019.

Darabi, Leila. Medical Question Answering for Persian. Master's thesis, LIACS, [۳۳] Leiden University, 2024.

Farahani, Mehrdad, et al. "Parsbert: Transformer-based model for persian language [۳۴] understanding." Neural Processing Letters 53 (2021): 3831-3847.

Jiang, Shuyang, et al. "MedS³: Towards Medical Slow Thinking with Self-Evolved [۳۵] Soft Dual-sided Process Supervision." arXiv preprint arXiv:2501.12051 .(2025)

Coulom, Rémi. "Efficient selectivity and backup operators in Monte-Carlo tree [۳۶] search." International conference on computers and games. Berlin, Heidelberg: Springer Berlin Heidelberg, .2006

Wu, Juncheng, et al. "Medreason: Eliciting factual medical reasoning steps in llms [۳۷] via knowledge graphs." arXiv preprint arXiv:2504.00993 .(2025)

Chen, Junying, et al. "Huatuogpt-o1, towards medical complex reasoning with [۳۸] llms." arXiv preprint arXiv:2412.18925 .(2024)

Guo, Daya, et al. "Deepseek-r1: Incentivizing reasoning capability in llms via rein- [۳۹] forcement learning." arXiv preprint arXiv:2501.12948 .(2025)

Wu, Tianhao, et al. "Thinking llms: General instruction following with thought gen- [۴۰] eration." arXiv preprint arXiv:2410.10630 .(2024)

Ho, Namgyu, Laura Schmid, and Se-Young Yun. "Large language models are rea- [۴۱] soning teachers." Proceedings of the 61st annual meeting of the association for com- putational linguistics (volume 1: long papers). 2023.

García-Ferrero, Iker, et al. "Medical mT5: an open-source multilingual text-to-text [۴۲] LLM for the medical domain." arXiv preprint arXiv:2404.07613 (2024).

Liu, Yang, et al. "Datasets for large language models: A comprehensive survey." [۴۳] arXiv preprint arXiv:2402.18041 (2024).

Yang, Songhua, et al. "Zhongjing: Enhancing the chinese medical capabilities of [۴۴] large language model through expert feedback and real-world multi-turn dialogue." Proceedings of the AAAI conference on artificial intelligence. Vol. 38. No. 17. 2024.

Bao, Zhijie, et al. "Disc-medllm: Bridging general large language models and real- [۴۵] world medical consultation." arXiv preprint arXiv:2308.14346 (2023).

Zhang, Hongbo, et al. "Huatuogpt, towards taming language model to be a doctor." [۴۶] arXiv preprint arXiv:2305.15075 (2023).

Wang, Xidong, et al. "Huatuo-26M, a Large-scale Chinese Medical QA Dataset." [۴۷] Findings of the Association for Computational Linguistics: NAACL 2025. 2025.

Zeng, Guangtao, et al. "MedDialog: Large-scale medical dialogue datasets." Pro- [۴۸] ceedings of the 2020 conference on empirical methods in natural language processing (EMNLP). 2020.

Han, Tianyu, et al. "MedAlpaca—an open-source collection of medical conversa- [۴۹] tional AI models and training data." arXiv preprint arXiv:2304.08247 (2023).

Bi, Xiao, et al. "Deepseek llm: Scaling open-source language models with longter- [۵۰] mism." arXiv preprint arXiv:2401.02954 .(2024)

Pal, Ankit, Logesh Kumar Umapathi, and Malaikannan Sankarasubbu. "Medmcqa: [۵۱] A large-scale multi-subject multi-choice dataset for medical domain question answer- ing." Conference on health, inference, and learning. PMLR, .2022

C-LARA-Instance, Manny Rayner ChatGPT. "How Woke is Grok? Empirical Evi- [۵۲] dence that xAI's Grok Aligns Closely with Other Frontier Models." .(2025)

Achiam, Josh, et al. "Gpt-4 technical report." arXiv preprint arXiv:2303.08774 [۵۳] .(2023)

Jamali, Naghmeh, et al. "PerMedCQA: Benchmarking Large Language Models [۵۴] on Medical Consumer Question Answering in Persian Language." arXiv preprint arXiv:2505.18331 .(2025)

Kalahroodi, Mohammad Javad Ranjbar, et al. "PersianMedQA: Language- [۵۵] Centric Evaluation of LLMs in the Persian Medical Domain." arXiv preprint arXiv:2506.00250 .(2025)

Yang, An, et al. "Qwen2 Technical Report." arXiv Preprint arXiv:2407.10671, 2024. [۵۶]

Dang, John, et al. "Aya expanse: Combining research breakthroughs for a new mul- [۵۷] tilingual frontier." arXiv preprint arXiv:2412.04261 (2024).

Team, Gemma, et al. "Gemma 2: Improving open language models at a practical [۵۸] size, 2024." URL <https://arxiv.org/abs/2408.00118> 1.3 (2024).

Rostami, Pedram, Ali Salemi, and Mohammad Javad Dousti. "Persian- [۵۹] mind: A cross-lingual persian-english large language model." arXiv preprint arXiv:2401.06466 (2024).

Dash, Saurabh, et al. "Aya Vision: Advancing the Frontier of Multilingual Multi- [۶۰] modality." arXiv preprint arXiv:2505.08751 (2025).

Shumailov, Ilia, et al. "AI models collapse when trained on recursively generated [۶۱] data." Nature 631.8022 (2024): 755-759.

Odumakinde, Ayomide, et al. "Multilingual arbitrage: Optimizing data pools to ac- [۶۲] celerate multilingual progress, 2024." URL <https://arxiv.org/abs/2408.14960>.

Hu, Edward J., et al. "Lora: Low-rank adaptation of large language models." ICLR [۶۳] 1.2 (2022): 3.

Dao, Tri. "Flashattention-2: Faster attention with better parallelism and work partitioning." arXiv preprint arXiv:2307.08691 (2023). [۶۴]

Lacoste, Alexandre, et al. "Quantifying the carbon emissions of machine learning." [۶۵] arXiv preprint arXiv:1910.09700 (2019).

Hurst, Aaron, et al. "Gpt-4o system card." arXiv preprint arXiv:2410.21276 (2024). [۶۶]

Khashabi, Daniel, et al. "Parsinlu: a suite of language understanding challenges for [۶۷] persian." Transactions of the Association for Computational Linguistics 9 (2021): 1147-1162.

Kashefi, Omid. "MIZAN: a large persian-english parallel corpus." arXiv preprint [۶۸] arXiv:1801.02107 (2018).

Bengio, Yoshua. "From system 1 deep learning to system 2 deep learning." Neural [۶۹] Information Processing Systems. .2019

Kahneman, Daniel. "Thinking, fast and slow." Farrar, Straus and Giroux .(2011) [۷۰]

Goyal, Anirudh, et al. "Recurrent independent mechanisms." arXiv preprint [۷۱] arXiv:1909.10893 .(2019)

Wang, Guan, et al. "Hierarchical Reasoning Model." arXiv preprint [۷۲] arXiv:2506.21734 .(2025)

Pan, Liangming, et al. "Automatically correcting large language models: Surveying [۷۳] the landscape of diverse self-correction strategies." arXiv preprint arXiv:2308.03188 .(2023)

Liang, Baoyu, Yuchen Wang, and Chao Tong. "AI Reasoning in Deep Learning Era: [۷۴] From Symbolic AI to Neural–Symbolic AI." Mathematics 13.11 (2025): .1707

Lightman, Hunter, et al. "Let's verify step by step." The Twelfth International Conference [۷۵] on Learning Representations. 2023.

Huang, Jie, and Kevin Chen-Chuan Chang. "Towards reasoning in large language [۷۶] models: A survey." Findings of the association for computational linguistics: ACL

2023. .2023

Shojaee, Parshin, et al. "The illusion of thinking: Understanding the strengths and [۷۷] limitations of reasoning models via the lens of problem complexity." arXiv preprint

arXiv:2506.06941 .(2025)

Wei, Jason, et al. "Emergent abilities of large language models." arXiv preprint [۷۸]

arXiv:2206.07682 .(2022)

Wei, Jason, et al. "Chain-of-thought prompting elicits reasoning in large language [۷۹]

models." Advances in neural information processing systems 35 (2022): 24824-

.24837

Brown, Bradley, et al. "Large language monkeys: Scaling inference compute with [۸۰]

repeated sampling." arXiv preprint arXiv:2407.21787 .(2024)

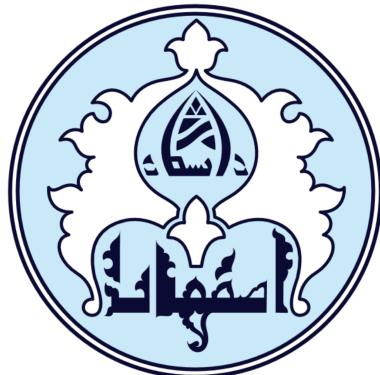
Zhang, Mozhi, et al. "Calibrating the confidence of large language models by eliciting [۸۱]

fidelity." arXiv preprint arXiv:2404.02655 .(2024)

Abstract

The use of artificial intelligence in answering medical questions is recognized as one of the emerging and critical fields in technology and healthcare, which has garnered widespread attention in recent years. This advanced technology, with its unique capabilities, can significantly enhance the quality of medical services provided to patients. Additionally, by accelerating the process of delivering medical information and providing quick and accurate responses to the questions of doctors and patients, it plays a vital role in reducing the workload of physicians. As such, artificial intelligence not only increases efficiency in healthcare systems but also improves the overall patient experience and paves the way for better and more effective treatments. On the other hand, since medicine is based on reasoning and logical analysis, developing a medical model designed on a chain of logical thoughts and reasoning can significantly enhance the accuracy and efficiency of such a model. This approach facilitates the execution of complex diagnostic and therapeutic processes in a more structured and purposeful manner. In this regard, every stage of diagnosis and treatment must be based on scientific evidence and reliable data. For instance, in the process of diagnosing diseases, doctors typically rely on medical history, clinical symptoms, and test results. By designing a logical model, these data can be interconnected in a logical chain, helping identify patterns and relationships between symptoms and diseases.

Keywords Artificial Intelligence in medicine, Persian language models, Medical language models, Natural language processing, Artificial Intelligence reasoning



**University of Isfahan
Faculty of Computer Engineering**

Developing a medical language model based on reasoning in Persian language

A Thesis submitted to the Graduate Studies Office
In partial fulfillment of the requirements for
The degree of Master of Science
in Computer Engineering - Artificial Intelligence and Robotics

By:

Mehrdad Ghassabi

Supervisor:

Dr. Hamidreza Baradaran Kashani

Dec 2025