

بسم الله الرحمن الرحيم



دانشگاه صنعتی اصفهان
دانشکده مهندسی برق و کامپیوتر

ارائه سازوکار فراموشی برای بهبود عملکرد تنظیم نظارت شده مدل های زبانی بزرگ

گزارش پروژه کارشناسی مهندسی کامپیوتر

علی ظاهری

استاد راهنما

دکتر عبدالرضا میرزائی

مهرماه ۱۴۰۴

تشر و قدردانی

پروردگار متعال را سپاسگزارم که با عطای صبر و پشتکار، مرا در انجام پروژه و گذراندن دوره کارشناسی یاری کرد. همچنین از دکتر میرزائی که با اعتماد و راهنمایی‌های دقیق و منظم خود، مسیر موفقیت را برایم هموار ساختند، صمیمانه قدردانی می‌کنم. بدون حمایت و همراهی ایشان، دستیابی به این دستاورد ممکن نمی‌شد. نقش ایشان در این موفقیت، برای من همواره ارزشمند و به‌یادماندنی خواهد بود.

فهرست مطالب

۹	چکیده
۱۰	فصل اول
۱۰	مقدمه
۱۱	۱-۱ هدف کلی و دستاوردها
۱۱	۱-۲ روش تحقیق
۱۲	۱-۳ ساختار گزارش
۱۳	فصل دوم
۱۳	کارهای مرتبط
۱۴	۱-۲ انتخاب داده برای تنظیم نظارت شده مدل های زبانی بزرگ
۱۵	۲-۲ یادزدایی و بهینه سازی ترجیحی در مدل های زبانی بزرگ
۱۶	۳-۲ تحلیل کبمودها و انگیزه پژوهش
۱۸	فصل سوم
۱۸	روش شناسی پژوهش
۱۹	۱-۳ مبانی نظری تنظیم نظارت شده
۱۹	۱-۳-۱ فرمولاسیون ریاضی تنظیم نظارت شده
۱۹	۱-۳-۲ تابع هدف استاندارد
۲۰	۱-۳-۳ مشکل کیفیت داده در تنظیم نظارت شده
۲۱	۱-۳-۴ محدودیت های روش های موجود
۲۱	۲-۳ محاسبه امتیاز کیفیت توکن ها
۲۱	۱-۲-۳ مفهوم تابع تأثیر در ارزیابی کیفیت
۲۲	۲-۲-۳ ایجاد مدل مرجع
۲۳	۳-۲-۳ فرمولاسیون تابع تأثیر سطح توکن
۲۴	۴-۲-۳ محاسبه امتیاز کیفیت
۲۵	۳-۳ الگوریتم تقسیم بندی توکن ها
۲۵	۱-۳-۳ اصول تقسیم بندی بر اساس امتیاز کیفیت
۲۶	۲-۳-۳ تعیین آستانه تقسیم بندی

۲۶.....	۳-۳-۳ فرآیند طبقه‌بندی توکن‌ها
۲۷.....	۴-۳-۳ ویژگی‌های الگوریتم تقسیم‌بندی
۲۸.....	۵-۳-۳ ملاحظات عملی و بهینه‌سازی
۲۸.....	۴-۳-۴ تابع هدف اصلاح‌شده و سازوکار فراموشی
۲۹.....	۱-۴-۳ مبانی نظری فراموشی انتخابی
۲۹.....	۲-۴-۳ طراحی تابع هدف دوگانه
۳۰.....	۳-۴-۳ سازوکار ضریب تعادل تطبیقی
۳۱.....	۴-۴-۳ پیاده‌سازی و جزئیات الگوریتم

۳۲ فصل چهارم

۳۲ نتایج و تحلیل یافته‌ها

۳۳.....	۱-۴ تنظیمات آزمایشی و پیکربندی محیط
۳۳.....	۴-۱-۱ انتخاب و تنظیم مجموعه داده‌های آموزشی
۳۴.....	۴-۱-۲ معیارهای ارزیابی و مجموعه داده‌های آزمون
۳۵.....	۴-۱-۳ انتخاب و پیکربندی مدل‌های آزمایشی
۳۶.....	۴-۱-۴ تنظیمات فنی و پیکربندی آموزش
۳۷.....	۲-۴ نتایج مقایسه‌ای و تحلیل عملکرد
۳۷.....	۴-۲-۱ مقایسه عملکرد بر روی مدل‌های مختلف اندازه
۳۹.....	۴-۲-۲ تحلیل حساسیت پارامترها
۴۱.....	۳-۲-۴ تأثیر نرخ فراموشی بر عملکرد
۴۲.....	۴-۲-۴ تحلیل الگوهای بهبود
۴۳.....	۴-۲-۵ پایداری و تکرارپذیری نتایج

۴۴ فصل پنجم

۴۴ نتیجه‌گیری و پیشنهادات برای تحقیقات آتی

۴۵.....	۱-۵ نتیجه‌گیری و دستاوردهای کلیدی
۴۵.....	۲-۵ محدودیت‌ها و چالش‌های موجود
۴۵.....	۳-۵ پیشنهادات برای تحقیقات آتی

۴۷ مراجع

فهرست اشکال

- شکل ۴-۱ تحلیل عملکرد روش فراموشی در مقابل نادیده گیری در مقادیر مختلف p ۳۹
- شکل ۴-۲ تحلیل عملکرد میانگین روش فراموشی با نرخ‌های مختلف فراموشی ۴۲

فهرست جداول

جدول ۴-۱	توزیع دقیق مجموعه داده‌های آموزشی	۳۳
جدول ۴-۲	مشخصات تفصیلی مجموعه داده‌های ارزیابی	۳۴
جدول ۴-۳	مقایسه عملکرد روش‌های مختلف بر روی مدل‌های LLaMA	۳۸
جدول ۴-۴	مقایسه عملکرد بر روی مدل LLaMA-2-13B	۳۹
جدول ۴-۵	تحلیل حساسیت پارامترهای t_{max} و t_{min}	۴۱

چکیده

تنظیم نظارت شده^۱ مدل‌های زبانی بزرگ نقش حیاتی در بهبود عملکرد این مدل‌ها برای کاربردهای تخصصی ایفا می‌کند. با این حال، کیفیت داده‌ها تأثیر مستقیمی بر موفقیت این فرآیند دارد و داده‌های نامناسب می‌توانند منجر به کاهش عملکرد یا حتی تخریب قابلیت‌های مدل شوند. روش‌های سنتی انتخاب داده با حذف کامل نمونه‌های نامناسب، حجم مجموعه داده‌ها را کاهش می‌دهند که این امر محدودیتی جدی محسوب می‌شود.

در این پژوهش، سازوکار جدیدی تحت عنوان "فراموشی" برای بهبود تنظیم نظارت شده مدل‌های زبانی بزرگ ارائه می‌شود. این روش به جای حذف کامل داده‌های نامناسب، توکن‌ها را به دو دسته مثبت و منفی تقسیم‌بندی کرده و برای توکن‌های منفی از تکنیک فراموشی استفاده می‌کند. بدین ترتیب، مدل نه تنها از توکن‌های مفید یاد می‌گیرد بلکه به صراحت توکن‌های گمراه‌کننده را فراموش می‌کند تا مرزهای دانش واضح‌تری تشکیل دهد.

آزمایش‌های انجام شده بر روی مدل‌های مختلف LLaMA و مجموعه داده‌های متنوع نشان می‌دهد که این روش عملکرد قابل توجهی نسبت به روش‌های سنتی دارد. نتایج حاکی از بهبود متوسط ۲/۵۱ تا ۸/۲۵ درصدی در عملکرد مدل‌ها بوده و همچنین تنوع بیشتری در پاسخ‌های تولیدی ایجاد می‌کند. این سازوکار با حفظ حجم کامل داده‌ها، راه‌حل مؤثری برای چالش تعادل بین کیفیت و کمیت داده‌ها ارائه می‌دهد [کد پیاده‌سازی].

^۱ Supervised Fine-tuning

فصل اول

مقدمه

در سال‌های اخیر، مدل‌های زبانی بزرگ^۱ به عنوان یکی از مهم‌ترین دستاوردهای هوش مصنوعی مطرح شده‌اند. این مدل‌ها که بر پایه معماری تبدیل‌گر^۲ ساخته شده‌اند، قابلیت‌های فوق‌العاده‌ای در درک و تولید متن طبیعی از خود نشان داده‌اند. با این حال، برای استفاده مؤثر از این مدل‌ها در کاربردهای تخصصی، فرآیند تنظیم نظارت‌شده ضروری است. تنظیم نظارت‌شده فرآیندی است که طی آن مدل‌های پیش‌آموزش‌دیده با استفاده از داده‌های کاربردی خاص، برای انجام وظایف تخصصی بهینه‌سازی می‌شوند. اما یکی از چالش‌های اساسی در این فرآیند، وابستگی شدید آن به کیفیت داده‌های آموزشی است. داده‌های نامناسب، نادرست یا گمراه‌کننده می‌توانند نه تنها مانع بهبود عملکرد شوند، بلکه باعث تنزل قابلیت‌های مدل نیز گردند.

^۱ Large Language Models

^۲ transformer

۱-۱ هدف کلی و دستاور

هدف اصلی این پژوهش، ارائه روشی نوآورانه برای حل مشکل کیفیت داده در تنظیم نظارت‌شده مدل‌های زبانی بزرگ است. برخلاف روش‌های سنتی که با حذف کامل داده‌های نامناسب، حجم مجموعه داده را کاهش می‌دهند، روش پیشنهادی بر پایه سازوکار فراموشی^۱ عمل می‌کند. این سازوکار با تقسیم‌بندی داده‌ها در سطح توکن به دو دسته مثبت و منفی، امکان استفاده بهینه از کل داده‌ها را فراهم می‌آورد. دستاورد اصلی این تحقیق، توسعه الگوریتمی است که قادر به شناسایی و فراموش کردن توکن‌های گمراه‌کننده بوده و در عین حال از توکن‌های مفید یادگیری می‌کند. این روش منجر به بهبود قابل توجه عملکرد مدل‌ها در کارهای مختلف و همچنین افزایش تنوع در پاسخ‌های تولیدی شده است.

۱-۲- روش تحقیق

در این پژوهش، برای بهبود دقت تنظیم نظارت‌شده مدل‌های زبانی بزرگ، از رویکردی چندمرحله‌ای استفاده شده است. ابتدا، یک مدل مرجع^۲ بر پایه بخشی از داده‌های آموزشی ایجاد شده است. سپس، با استفاده از تابع تأثیر^۳ بین مدل پایه و مدل مرجع، کیفیت هر توکن ارزیابی شده است. مرحله بعد، تقسیم‌بندی توکن‌ها بر اساس امتیاز کیفیت محاسبه‌شده انجام می‌شود. توکن‌هایی که امتیاز مثبت دارند، به عنوان توکن‌های مفید طبقه‌بندی شده و توکن‌هایی با امتیاز منفی، به عنوان توکن‌های نامناسب در نظر گرفته می‌شوند. در ادامه، تابع هدف آموزش به گونه‌ای تعدیل می‌شود که برای توکن‌های مثبت از روش معمول حداکثرسازی درستنمایی^۴ استفاده شود، اما برای توکن‌های منفی، احتمال تولید آن‌ها به صراحت کاهش یابد.

برای ارزیابی روش پیشنهادی، آزمایش‌های گسترده‌ای بر روی مدل‌های مختلف LLaMA شامل LLaMA 3.2 1B، LLaMA 3.2 3B، LLaMA 3.1 8B و LLaMA 2 13B انجام شده است. مجموعه داده‌های آموزشی^۵ شامل ترکیبی از پنج مجموعه داده معتبر Flan_v2، Dolly، OpenAssistant، Stanford Alpaca و WizardLM است. ارزیابی عملکرد بر پایه پنج معیار سنجش^۶ TruthfulQA، BoolQ، LogiQA، TydiQA و ASDiv صورت گرفته است.

^۱ Forgetting

^۲ Reference Model

^۳ Influence Function

^۴ Maximum Likelihood

^۵ Train dataset

^۶ Benchmark

۱-۳- ساختار گزارش

فصل دوم، به بررسی کامل و دقیق از مطالعات پیشین در زمینه تنظیم نظارت شده مدل‌های زبانی بزرگ و روش‌های انتخاب داده می‌پردازد. این فصل شامل تحلیل روش‌های سنتی انتخاب داده، محدودیت‌های آن‌ها، و همچنین مروری بر تکنیک‌های یادزدایی^۱ و بهینه‌سازی ترجیحی^۲ است.

فصل سوم، به معرفی و توضیح روش‌شناسی پیشنهادی اختصاص دارد. در این فصل، مبانی نظری سازوکار فراموشی، نحوه محاسبه امتیاز کیفیت توکن‌ها، الگوریتم تقسیم‌بندی، و تابع هدف اصلاح شده به تفصیل ارائه می‌شود.

فصل چهارم، نتایج حاصل از پیاده‌سازی و تجربیات انجام شده را به همراه تحلیل و بررسی آن‌ها ارائه می‌دهد. مقایسه عملکرد روش پیشنهادی با روش‌های پایه، تحلیل حساسیت پارامترها، و بررسی تأثیر نسبت توکن‌های مثبت و منفی در این فصل مطرح می‌شود.

فصل پنجم، به نتیجه‌گیری و ارائه پیشنهادات برای تحقیقات آتی اختصاص دارد. هدف از این فصل، ارائه یک چهارچوب جامع برای تعیین مسیرهای آتی پژوهش و تعیین جایگاه کار انجام شده در حوزه مدل‌های زبانی بزرگ است.

^۱ Unlearning

^۲ Preference optimization

فصل دوم

کارهای مرتبط

این فصل به بررسی مطالعات پیشین در زمینه تنظیم نظارت شده مدل‌های زبانی بزرگ و تمرکز آن بر مشکل کیفیت داده‌ها و راه‌حل‌های ارائه شده در این حوزه می‌پردازد. با توجه به اهمیت روزافزون مدل‌های زبانی بزرگ در کاربردهای مختلف و نیاز به بهینه‌سازی عملکرد آن‌ها برای وظایف تخصصی، بررسی روش‌های انتخاب داده و تکنیک‌های بهبود فرآیند تنظیم ضروری است. ابتدا، روش‌های سنتی انتخاب داده و محدودیت‌های آن‌ها در حفظ مقیاس مجموعه داده بررسی می‌شود. سپس، به تکنیک‌های یادزدایی و بهینه‌سازی ترجیحی پرداخته می‌شود که راه‌حلی برای مواجهه با داده‌های نامناسب ارائه می‌دهند. در نهایت، کمبودهای موجود در ادبیات و انگیزه‌های اصلی برای توسعه سازوکار فراموشی، به‌ویژه در حفظ مقیاس کامل داده‌ها، مورد تحلیل قرار خواهد گرفت.

۲-۱- انتخاب داده برای تنظیم نظارت شده مدل های زبانی بزرگ

تنظیم نظارت شده به عنوان یکی از مهم ترین تکنیک های بهبود عملکرد مدل های زبانی بزرگ، بر اساس اصل حداکثرسازی درستمایی عمل می کند و فرض می کند که تمامی توکن های موجود در مجموعه داده حاوی اطلاعات مفید هستند [۱]. اما تحقیقات نشان داده اند که کیفیت داده ها نقش تعیین کننده ای در موفقیت این فرآیند دارد و خطاهای ناشی از منابع مختلف نظیر حاشیه نویسی های انسانی^۱، ابزارهای خودکار حاشیه نویسی^۲، توهامات مدل های زبانی بزرگ^۳ و عدم تطابق در پردازش داده ها می تواند عملکرد نهایی را تحت تأثیر قرار دهد [۲].

مطالعه LIMA در [۳] فرضیه ای را مطرح کرد مبنی بر اینکه مدل های زبانی بزرگ عمدتاً سبک پاسخ دهی مجموعه داده را یاد می گیرند تا اینکه دانش پیش آموزش دیده خود را به سمت وظایف تخصصی به روزرسانی کنند. این مطالعه نشان داد که تنظیم نظارت شده بر روی ۱۰ هزار نمونه به دقت انتخاب شده، عملکرد بهتری نسبت به استفاده از مجموعه داده های بزرگ تر ارائه می دهد.

برای مواجهه با چالش های کیفی، پژوهشگران مزایای کیفیت داده نسبت به کمیت آن را بررسی کرده و الگوریتم های انتخاب مبتنی بر معیارهای کیفیت و تنوع را برای فیلتر کردن نمونه های گمراه کننده و بهبود قابلیت های پیروی از دستورالعمل پیشنهاد داده اند [۴، ۵، ۶، ۷]. اگرچه این روش ها در بهبود عملکرد مؤثر هستند، اما محدودیت اساسی دارند: آن ها در سطح نمونه عمل می کنند و با دور انداختن کامل مثال ها، مقیاس کلی داده های موجود برای آموزش را کاهش می دهند.

معیارهای مختلفی برای ارزیابی کیفیت داده معرفی شده اند از جمله تطبیق گرایان [۸]، بازخورد انسانی [۹] و امتیازهای تابع تأثیر^۴ [۱۰]. علاوه بر این، در [۱۰] نشان دادند که امتیازات تأثیر طبیعی بالاتر برای وظایف خاص می تواند سوگیری^۵ در انتخاب داده ایجاد کند و پیشنهاد دادند که امتیازات تأثیر را در میان وظایف مختلف نرمال سازی کرده و سپس به صورت تکراری نمونه ها را برای مهارت های کم نمایندگی انتخاب کنند.

در رویکرد دیگری [۱۱] چهارچوب دو مرحله ای مقاوم در برابر نویز پیشنهاد دادند که تشخیص نویز را با استفاده از سیستم های متخصص چند گانه انجام می دهد و سپس داده های وظیفه پایین دستی را با یافتن مثال های مشابه از مجموعه تمیز برای ارائه زمینه، برچسب گذاری مجدد می کند. همچنین محققان نشان دادند که انتخاب نمونه های آموزشی منطبق با دانش موجود مدل می تواند عملکرد را از طریق تولید جفت های دستورالعمل-پاسخ چند گانه و انتخاب آن هایی با

¹ Human annotators

² Automatic annotation tools

³ LLM hallucinations

⁴ Influence function scores

⁵ Bias

بالاترین احتمال مطابق با مدل هدف بهبود دهد [۱۲].

مطالعات اخیر الگوریتم‌های انتخاب داده با کیفیت بالا برای تنظیم نظارت‌شده مدل‌های زبانی بزرگ را بررسی کرده‌اند، اما نکته مهمی را نادیده گرفته‌اند: حتی در نمونه‌های پرنویز، برخی توکن‌ها همچنان حاوی اطلاعات ارزشمند هستند. با دور انداختن کامل نمونه‌ها، این روش‌ها ناخواسته سیگنال‌های آموزشی مفید را حذف می‌کنند. علاوه بر این، این رویکردها در استفاده از داده‌های رد شده به عنوان سیگنال یادگیری ناکام هستند.

۲-۲ یادزدایی و بهینه‌سازی ترجیحی در مدل‌های زبانی بزرگ

چندین رویکرد برای حذف اطلاعات خاص از مدل‌های زبانی بزرگ بدون بازآموزی^۱ کامل آن‌ها از ابتدا پیشنهاد شده است. این روش‌ها شامل استراتژی‌های جایگزینی و برچسب‌گذاری مجدد داده [۱۳، ۱۴] و تکنیک‌های ویرایش دانش از طریق پیش‌بینی به‌روزرسانی‌های هدفمند پارامتر برای تغییر حقایق خاص ضمن حفظ سایر دانش‌ها [۱۵] می‌شود.

روش‌های مبتنی بر صعود گرادیان^۲ معمولاً به دلیل سادگی خود استفاده می‌شوند که در ستنمایی منفی لگاریتمی^۳ توالی‌های توکن خاص را حداکثر می‌کنند [۱۶، ۱۷، ۱۸، ۱۹، ۲۰]. با این حال، برخی از آن‌ها منجر به تنزل در خروجی‌های مدل‌های زبانی بزرگ به صورت کلی شده و تمامیت کلی مدل‌ها را هنگام حذف دانش هدف آسیب می‌رسانند [۲۱، ۲۲، ۲۳، ۲۴، ۲۵] که یادزدایی مفرط^۴ نامیده می‌شود. برای مقابله با این مشکل، تکنیک‌های تنظیم‌سازی نظیر کمینه‌سازی واگرایی KL^۵ بین توزیع‌های خروجی مدل‌های پیش‌آموزش‌دیده و تنظیم‌شده پیشنهاد شده است [۲۶]. این امر محاسبات اضافی و حساسیت ابرپارامتر^۶ ایجاد می‌کند.

همچنین در [۲۷] روش صعود وزن‌دار گرادیان^۷ را معرفی کردند که وزن‌های مبتنی بر اعتماد را برای کاهش یادزدایی مفرط به صورت کنترل‌شده اعمال می‌کند. در حوزه بهینه‌سازی ترجیحی، بهینه‌سازی مستقیم ترجیحات^۸ به عنوان جایگزینی برای روش‌های مبتنی بر بهینه‌سازی سیاست نزدیک^۹ ظهور کرده است. اگرچه بهینه‌سازی سیاست نزدیک به دلیل کارایی نمونه در مقایسه با روش‌های گرادیان سیاست پیشین موفق بوده، اما همچنان از مدل‌سازی صریح مدل پاداش و

^۱ Retraining

^۲ Gradient Ascent

^۳ Negative log-likelihood

^۴ Excessive unlearning

^۵ KL divergence

^۶ Hyperparameter

^۷ Weighted Gradient Ascent (WGA)

^۸ Direct Preference Optimization (DPO)

^۹ Proximal Policy Optimization (PPO)

تنظیم پیچیده ابرپارامتر رنج می‌برد [۲۸]. برای مقابله با این چالش‌ها و مقاوم‌تر و کم‌هزینه‌تر کردن محاسبات، بهینه‌سازی مستقیم ترجیحات برای هدف تراز^۱ را به فرمول‌بندی حداکثر درستمایی بر روی داده‌های جفت‌شده ترجیحی تبدیل می‌کند و سعی می‌کند پاسخ‌های ترجیح‌داده‌شده را محتمل‌تر و پاسخ‌های غیرترجیحی را کم‌احتمال‌تر کند [۲۹]. مطالعات گسترده‌ای برای رفع محدودیت‌های بهینه‌سازی مستقیم ترجیحات انجام شده است [۳۰، ۳۱، ۳۲، ۳۳، ۳۴، ۳۵]. رویکرد جدیدی برای یادزدایی مبتنی بر ترجیحات^۲ در [۱۷] پیشنهاد شد که مجموعه فراموشی را به عنوان پاسخ‌های غیرترجیحی تعریف می‌کند و پاسخ ترجیحی شامل پاسخ‌های انکاری نظیر "من پاسخ را نمی‌دانم" است. با الهام از این تحقیق، در [۳۶] واریانت جدیدی از بهینه‌سازی مستقیم ترجیحات به نام بهینه‌سازی ترجیح منفی^۳ پیشنهاد دادند که تنها از پاسخ‌های منفی استفاده می‌کند و پاسخ‌های مثبت را نادیده می‌گیرد. همچنین در [۲۷] بهینه‌سازی ترجیح منفی سطح توکن^۴ و نسخه وزن‌دار^۵ آن را پیشنهاد دادند که یادزدایی را در سطح توکن فردی برای کنترل دقیق‌تر بر حذف دانش اعمال می‌کند، اما این روش‌ها به طور خاص برای فراموشی هدفمند توسعه یافته‌اند تا اینکه مکمل یادگیری در طول تنظیم نظارت‌شده باشند.

۲-۳- تحلیل کمبودها و انگیزه پژوهش

بررسی ادبیات موجود نشان می‌دهد که اگرچه پیشرفت‌های قابل توجهی در زمینه انتخاب داده و یادزدایی صورت گرفته، اما کمبودهای مهمی در این حوزه وجود دارد. روش‌های سنتی انتخاب داده با عملکرد در سطح نمونه، مجبور به دور انداختن کامل مثال‌هایی هستند که ممکن است حاوی اطلاعات ارزشمند باشند. به عنوان مثال، یک جفت دستورالعمل-پاسخ ممکن است دارای دستورالعمل مناسب و قسمت‌هایی از پاسخ صحیح باشد، اما به دلیل وجود توکن‌های نامناسب در بخش‌هایی از پاسخ، کل نمونه حذف می‌شود. این امر تعادل اجتناب‌ناپذیری بین کیفیت و کمیت ایجاد می‌کند که همچنان حل نشده باقی مانده است و منجر به هدر رفتن منابع محاسباتی و اطلاعات مفید می‌شود.

از سوی دیگر، تکنیک‌های یادزدایی موجود عمدتاً برای حذف اطلاعات خاص از مدل‌های آموزش‌دیده طراحی شده‌اند؛ مانند حذف اطلاعات کپی‌رایت یا محتوای نامناسب، تا اینکه به عنوان سازوکار بهبود کلی فرآیند تنظیم نظارت‌شده عمل کنند. این روش‌ها معمولاً در مراحل پس از آموزش اولیه اعمال می‌شوند و هدف اصلی آن‌ها حذف دانش خاص است، نه بهبود یادگیری از داده‌های جدید. همچنین، اکثر این روش‌ها نیاز به تنظیم دقیق پارامترها دارند،

¹ Alignment objective

² preference-based unlearning

³ Negative Preference Optimization (NPO)

⁴ Token-level Negative Preference Optimization (TNPO)

⁵ Weighted Token-level Negative Preference Optimization (WTNPO)

حساسیت بالایی نسبت به انتخاب ابرپارامترها نشان می‌دهند و ممکن است منجر به یادزدایی مفرط شوند که قابلیت‌های کلی مدل را تضعیف کند.

علاوه بر این، روش‌های موجود انتخاب داده اغلب بر معیارهای سطح نمونه متکی هستند که نمی‌توانند تمایز دقیقی بین بخش‌های مفید و نامناسب درون یک نمونه واحد ایجاد کنند. این محدودیت باعث می‌شود که یا کل نمونه حفظ شود و ممکن است شامل توکن‌های گمراه‌کننده باشد، یا کل آن حذف گردد و منجر به از دست رفتن اطلاعات مفید شود. نوآوری اصلی این پژوهش [۳۷] در ترکیب مزایای هر دو رویکرد است: حفظ مقیاس کامل داده‌ها همانند روش‌های سنتی تنظیم نظارت‌شده، در حالی که از سازوکار فراموشی برای مواجهه با توکن‌های نامناسب استفاده می‌کند. این رویکرد نوآورانه عملکرد در سطح توکن را امکان‌پذیر می‌سازد، به طوری که هر توکن بر اساس کیفیت و مفید بودن آن ارزیابی شده و تصمیم‌گیری مناسب برای آن اتخاذ می‌شود.

بدین ترتیب، مدل قادر خواهد بود تا از توکن‌های مفید یادگیری کند و در عین حال توکن‌های گمراه‌کننده را به صراحت فراموش نماید. این فرآیند دوگانه منجر به تشکیل مرزهای دانشی واضح‌تری می‌شود، به طوری که مدل نه تنها می‌داند چه چیزی را باید یاد بگیرد، بلکه چه چیزی را نباید تولید کند نیز درک می‌کند. این سازوکار همچنین امکان استفاده بهینه از تمام داده‌های موجود را فراهم می‌آورد، زیرا هیچ نمونه‌ای به طور کامل حذف نمی‌شود و در عوض، هر نمونه بر اساس محتوای توکن‌های آن به صورت انتخابی در فرآیند یادگیری یا فراموشی شرکت می‌کند.

علاوه بر این، این روش چالش تعادل بین کیفیت و کمیت را نیز حل می‌کند، چرا که دیگر نیازی به انتخاب بین حفظ یا حذف کامل نمونه‌ها نیست. در نهایت، این سازوکار امکان تعمیم‌پذیری بهتر مدل را فراهم می‌آورد، زیرا مدل یاد می‌گیرد که چگونه الگوهای مناسب را تشخیص دهد و از الگوهای نامناسب اجتناب کند، که این مهارت در مواجهه با داده‌های جدید نیز قابل استفاده است.

فصل سوم

روش‌شناسی پژوهش

این فصل به تشریح کامل روش‌شناسی سازوکار فراموشی برای بهبود تنظیم نظارت‌شده مدل‌های زبانی بزرگ می‌پردازد. هسته اصلی روش پیشنهادی بر پایه تقسیم‌بندی هوشمند توکن‌ها به دو دسته مثبت و منفی و اعمال استراتژی‌های متفاوت یادگیری برای هر دسته استوار است.

ساختار این فصل شامل چهار بخش اصلی است: ابتدا مبانی نظری و فرمولاسیون ریاضی تنظیم نظارت‌شده سنتی و چالش‌های آن ارائه می‌شود. سپس، روش محاسبه امتیاز کیفیت توکن‌ها با استفاده از تابع تأثیر و مقایسه مدل پایه با مدل مرجع تشریح می‌گردد. در ادامه، الگوریتم تقسیم‌بندی توکن‌ها بر اساس آستانه‌گذاری بر روی امتیازات کیفیت توضیح داده می‌شود. بخش پایانی، طراحی تابع هدف ترکیبی و نحوه تعادل بین اجزای یادگیری و فراموشی با استفاده از ضریب تطبیقی را شرح می‌دهد.

تمامی مراحل الگوریتم به همراه جزئیات پیاده‌سازی، تنظیمات ابرپارامترها و ملاحظات عملی اجرای روش به تفصیل مستند شده است.

۳-۱- مبانی نظری تنظیم نظارت شده

۳-۱-۱- فرمولاسیون ریاضی تنظیم نظارت شده

تنظیم نظارت شده در مدل های زبانی بزرگ بر پایه مدل سازی زبان خودهمبسته^۱ عمل می کند که به عنوان پیش بینی ترتیبی خروجی ها بر اساس زمینه قبلی تعریف می شود. این فرآیند نقش غالبی در مدل های زبانی معاصر ایفا می کند. پس از پیش آموزش، تنظیم نظارت شده معمولاً برای بهبود بیشتر مدل های زبانی بزرگ در وظایف خاص با بهینه سازی بر روی جفت های دستورالعمل-پاسخ خاص وظیفه اتخاذ می شود.

مجموعه آموزشی شامل جفت های توالی نمونه است که هر جفت حاوی پرامپت ورودی و پاسخ تکمیلی می باشد. هر پرامپت دارای طول توالی مشخص بوده و به طور مشابه، هر پاسخ تکمیلی نیز دارای طول توالی خاص خود است:

$$D = \{(X_i, Y_i)\}_{i=1}^N \quad (۱-۳)$$

که در آن N تعداد کل جفت های نمونه را نشان می دهد. در رویکرد خودهمبسته، مدل یاد می گیرد که توزیع احتمال را برای هر توکن در پاسخ تخمین بزند. این توزیع بر اساس کل پرامپت و تمام توکن های تولید شده قبلی در پاسخ شرطی است:

$$P(y_{i,j} | X_i, y_{i,:j}; \theta) \quad (۲-۳)$$

که در آن $y_{i,j}$ نشان دهنده j امین توکن از i امین پاسخ، X_i پرامپت ورودی، $y_{i,:j}$ تمام توکن های تولید شده قبلی در پاسخ شرطی است، و θ پارامترهای مدل را نشان می دهد.

۳-۱-۲- تابع هدف استاندارد

هدف متداول آنتروپی متقاطع^۲ معمولاً برای تنظیم نظارت شده اتخاذ می شود. این تابع هدف بر اساس اصل حداکثر سازی درستی شرطی عمل می کند و فرض می کند که تمامی توکن های موجود در مجموعه داده حاوی اطلاعات مفید و

^۱ Autoregressive

^۲ cross-entropy

قابل اعتماد هستند. در این رویکرد، وزن یکسانی برای تمام توکن‌ها در نظر گرفته می‌شود و مدل سعی می‌کند احتمال تولید تمام توکن‌های هدف را حداکثر کند:

$$\mathcal{L}(\theta) = \frac{1}{\sum_{(i,j) \in \mathcal{I}} w_{i,j}} \sum_{(i,j) \in \mathcal{I}} -\log P(y_{i,j} | X_i, y_{i,:j}; \theta) \quad (3-3)$$

که در آن مجموعه شاخص به صورت زیر تعریف می‌شود:

$$\mathcal{I} := \{(i,j) | i \in 1,2, \dots, N, j \in 1,2, \dots, n_i\}$$

و تابع زیان^۱ هر توکن نیز به صورت منفی لگاریتم احتمال آن توکن محاسبه می‌گردد:

$$l(y_{i,j} | X_i, y_{i,:j}; \theta) := -\log P(y_{i,j} | X_i, y_{i,:j}; \theta)$$

۳-۱-۳ مشکل کیفیت داده در تنظیم نظارت‌شده

مدل‌های زبانی بزرگ قابلیت‌ها و بازنمایی‌های دانشی متنوعی را از طریق پیش‌آموزش بر روی مجموعه داده‌های گسترده کسب می‌کنند. با این حال، برای استفاده از آن‌ها در وظایف تخصصی، تکنیک‌هایی مانند تنظیم نظارت‌شده نقش قابل توجهی در تقویت عملکرد آن‌ها ایفا می‌کند. این فرآیند معمولاً بدون هیچ انتخاب یا حذف در اجزای مجموعه داده انجام می‌شود.

جمع‌آوری داده‌های با کیفیت بالا که دانش خاص مورد نیاز را نمایندگی کند، برای جلوگیری از عدم دقت و تراز مؤثر مدل زبانی بزرگ حیاتی است. این فرآیند در عمل به دلیل عوامل مختلفی می‌تواند چالش‌برانگیز باشد. داده‌های خاص وظیفه اغلب کمیاب هستند، به ویژه در حوزه‌های نوظهور. علاوه بر این، مجموعه داده‌ها از منابع مختلف جمع‌آوری می‌شوند که اغلب منجر به سبک‌های زبانی و کیفیت ناسازگار می‌شود.

خطاهای ناشی از استفاده از ابزارهای حاشیه‌نویس، حاشیه‌نویسی دستی انسان، توهّمات مدل‌های زبانی بزرگ و عدم تطابق در پردازش داده‌ها نیز مشکلات اضافی ایجاد می‌کند. هر یک از این عوامل می‌تواند توکن‌های پرنویز و

^۱ Loss Function

گمراه‌کننده‌ای را به مجموعه داده اضافه کند که فرآیند بهینه‌سازی را به خطر انداخته و منجر به تعمیم ضعیف شود.

۳-۱-۴- محدودیت‌های روش‌های موجود

روش‌های موجود برای کاهش تأثیرات داده‌های کم کیفیت و گمراه‌کننده، روش‌های مختلف انتخاب داده را پیشنهاد داده‌اند تا داده‌های مفید و با کیفیت بالا را برای تنظیم حفظ کنند. این روش‌ها عمدتاً فیلترینگ داده را در سطح نمونه انجام می‌دهند، در حالی که فیلترینگ سطح توکن می‌تواند مقیاس مجموعه داده و اطلاعات ریزدانه را بسیار بیشتر حفظ کند.

محدودیت اساسی روش‌های سنتی این است که آن‌ها داده‌های کم کیفیت را در طول تنظیم به طور کامل دور می‌اندازند. این کار به طور قابل توجهی مقیاس مجموعه داده اصلی را کاهش می‌دهد و به طور بالقوه تعمیم مدل را محدود می‌کند. علاوه بر این، این رویکردها قادر به استفاده از اطلاعات موجود در داده‌های رد شده نیستند.

سؤال اساسی که باقی می‌ماند این است: چگونه می‌توان از کل مجموعه داده آموزشی در مقیاس اصلی آن استفاده کرد و در عین حال عملکرد مدل را بهبود بخشید؟ آیا ممکن است نه تنها از نمونه‌های با کیفیت بالا یاد بگیریم بلکه از داده‌ها و توکن‌های گمراه‌کننده برای ایجاد مرزهای دانشی واضح‌تر بدون بیش‌برازش^۱ به نوز استفاده کنیم؟ چنین رویکردی می‌تواند منجر به بهبود در تعمیم مدل شود و در عین حال دامنه جامع مجموعه داده اصلی را حفظ کند.

۳-۲- محاسبه امتیاز کیفیت توکن‌ها

۳-۲-۱- مفهوم تابع تأثیر در ارزیابی کیفیت

برای کمی‌سازی کیفیت توکن، از مفهوم تابع تأثیر بین مدل پایه و مدل مرجع استفاده می‌شود. تابع تأثیر در اصل ابزاری قدرتمند برای درک و تحلیل تأثیر نقاط داده فردی بر عملکرد کلی مدل‌های یادگیری ماشین است. این مفهوم که ریشه در نظریه اقتصاد و آمار دارد، امکان شناسایی و ارزیابی کمی نمونه‌هایی را فراهم می‌آورد که بیشترین تأثیر مثبت یا منفی را بر روی فرآیند یادگیری و عملکرد نهایی مدل اعمال می‌کنند.

در زمینه مدل‌های زبانی بزرگ، تابع تأثیر می‌تواند به شناسایی نمونه‌های آموزشی کمک کند که حذف یا اضافه کردن آن‌ها تأثیر قابل توجهی بر رفتار مدل خواهد داشت. با این حال، یکی از نوآوری‌های کلیدی این پژوهش، گسترش این مفهوم از سطح نمونه به سطح توکن است. این گسترش امکان ارزیابی دقیق‌تر و ریزدانه‌تر کیفیت داده‌ها را فراهم

^۱ Overfitting

می‌آورد، زیرا در عمل، یک نمونه آموزشی ممکن است شامل ترکیبی از توکن‌های مفید و نامناسب باشد.

ایده بنیادی که در پشت این رویکرد قرار دارد، بر این اصل استوار است که توکن‌هایی که پس از یک دوره آموزش اولیه قابلیت پیش‌بینی بهتری پیدا کرده‌اند، نمایانگر الگوها و ساختارهای زبانی معتبری هستند که مدل توانسته به طور موفقیت‌آمیز درک کرده و از آن‌ها یاد بگیرد. این توکن‌ها احتمالاً حاوی اطلاعات مفید، ساختارهای گرامری صحیح، یا محتوای معنایی ارزشمندی می‌باشند که برای بهبود عملکرد مدل مفید است.

در مقابل، توکن‌هایی که حتی پس از آموزش همچنان پیش‌بینی دشواری دارند، یا به طور غیرمنتظره‌ای قابلیت پیش‌بینی آن‌ها در مدل آموزش دیده کاهش یافته است، ممکن است نشان‌دهنده مشکلاتی در کیفیت داده باشند. این مشکلات می‌تواند شامل نویز در داده‌ها، خطاهای تایپی، ساختارهای گرامری نادرست، محتوای متناقض، یا حتی اطلاعات کاملاً گمراه‌کننده باشد که نه تنها برای یادگیری مدل مفید نیست، بلکه ممکن است تأثیر منفی بر عملکرد کلی آن داشته باشد.

۳-۲-۲- ایجاد مدل مرجع

گام اول و بنیادی در محاسبه امتیاز کیفیت توکن‌ها، ایجاد مدل مرجع است که به عنوان نقطه مقایسه برای ارزیابی تأثیر هر توکن عمل می‌کند. این مدل مرجع از طریق اعمال فرآیند تنظیم نظارت‌شده بر روی مدل پایه و با استفاده از زیرمجموعه‌ای دقیقاً انتخاب‌شده از داده‌های آموزشی ایجاد می‌شود. انتخاب این زیرمجموعه یکی از مراحل حساس و مهم فرآیند است که باید با دقت و توجه به اصول آماری صورت گیرد.

معیارهای انتخاب زیرمجموعه مرجع شامل چندین اصل مهم است. نخست، این زیرمجموعه باید از نظر آماری نماینده کل مجموعه داده باشد تا اطمینان حاصل شود که مدل مرجع توانایی تعمیم‌پذیری مناسبی دارد و الگوهای کلی موجود در داده‌ها را به درستی یاد گرفته است. دوم، اندازه این زیرمجموعه باید به گونه‌ای انتخاب شود که از یک سو کافی برای آموزش مؤثر مدل باشد و از سوی دیگر از نظر محاسباتی مقرون به صرفه و عملی باقی بماند.

فرض کنید $D_{ref} \subset D$ زیرمجموعه مرجع را نمایش دهد که شرط اساسی $D_{ref} \cap D_{train} = \emptyset$ را برآورده می‌کند. این شرط عدم تداخل از اهمیت بالایی برخوردار است زیرا اطمینان می‌دهد که مقایسه بین مدل پایه و مدل مرجع بر اساس داده‌های کاملاً مستقل و غیرمتعارض صورت گیرد. نقض این شرط می‌تواند منجر به سوگیری در ارزیابی کیفیت توکن‌ها شود و نتایج غیرقابل اعتمادی ایجاد کند. مدل مرجع θ' از طریق حل مسئله بهینه‌سازی زیر به دست می‌آید:

$$\theta' = \arg \min_{\theta} \mathcal{L}(\theta; D_{ref}) \quad (۴-۳)$$

که در آن $\mathcal{L}(\theta; D_{ref})$ تابع هدف استاندارد آنتروپی متقاطع محاسبه شده بر روی زیرمجموعه مرجع است. این فرآیند بهینه‌سازی معمولاً با استفاده از الگوریتم‌های گرادیان نزولی و تکنیک‌های بهینه‌سازی مدرن مانند Adam یا AdamW انجام می‌شود.

یکی از مزایای مهم این رویکرد این است که مدل مرجع نمایانگر حالتی از مدل است که بر روی زیرمجموعه‌ای از داده‌ها آموزش دیده و توانسته الگوهای قابل یادگیری موجود در آن زیرمجموعه را درک کند. بنابراین، مقایسه عملکرد این مدل با مدل پایه می‌تواند اطلاعات ارزشمندی در مورد قابلیت یادگیری هر توکن فراهم آورد.

۳-۲-۳ فرمولاسیون تابع تأثیر سطح توکن

پس از ایجاد مدل مرجع، گام بعدی محاسبه تابع تأثیر برای هر توکن است. این تابع معیاری کمی از تأثیر هر توکن بر فرآیند یادگیری ارائه می‌دهد و بر اساس مقایسه عملکرد مدل پایه و مدل مرجع در پیش‌بینی آن توکن محاسبه می‌شود. فرمولاسیون ریاضی این تابع به صورت زیر تعریف می‌گردد:

$$\text{Inf}(y_{i,j} | X_i, y_{i,:j}; \theta, \theta') = l(y_{i,j} | X_i, y_{i,:j}; \theta') - l(y_{i,j} | X_i, y_{i,:j}; \theta) \quad (۵-۳)$$

در این فرمول، $l(y_{i,j} | X_i, y_{i,:j}; \theta)$ نشان‌دهنده زیان محاسبه شده توسط مدل پایه برای توکن $y_{i,j}$ است، در حالی که $l(y_{i,j} | X_i, y_{i,:j}; \theta')$ همین زیان را برای مدل مرجع نمایش می‌دهد. تفاوت این دو مقدار، تأثیر آموزش بر قابلیت پیش‌بینی آن توکن خاص را نشان می‌دهد.

تفسیر این تابع از اهمیت ویژه‌ای برخوردار است. مقدار مثبت تابع تأثیر به این معناست که زیان توکن در مدل مرجع بیشتر از مدل پایه است، که نشان‌دهنده کاهش قابلیت پیش‌بینی آن توکن پس از آموزش است. این وضعیت ممکن است نشان‌دهنده این باشد که توکن مذکور دارای ساختار پیچیده‌ای است که مدل هنوز نتوانسته به طور کامل یاد بگیرد، یا اینکه حاوی نویز یا اطلاعات متناقضی است که فرآیند یادگیری را مختل می‌کند.

برعکس، مقدار منفی تابع تأثیر نشان‌دهنده کاهش زیان در مدل مرجع نسبت به مدل پایه است، که به معنای بهبود قابلیت پیش‌بینی آن توکن پس از آموزش اولیه می‌باشد. این وضعیت نشان می‌دهد که توکن مذکور حاوی الگوها یا اطلاعاتی

است که مدل توانسته به طور موفقیت آمیز از آن یاد بگیرد و قابلیت پیش‌بینی خود را در مورد آن بهبود دهد. مقدار صفر یا نزدیک به صفر تابع تأثیر نیز حالت جالب توجهی است که نشان می‌دهد آموزش تأثیر قابل توجهی بر قابلیت پیش‌بینی آن توکن نداشته است. این وضعیت می‌تواند نشان‌دهنده توکن‌هایی باشد که یا بسیار ساده و قابل پیش‌بینی هستند که نیاز به آموزش اضافی ندارند، یا دارای پیچیدگی متوسطی هستند که نه مفید و نه مضر محسوب می‌شوند.

۳-۲-۴- محاسبه امتیاز کیفیت

گام نهایی در این فرآیند، تبدیل تابع تأثیر به امتیاز کیفیت است که به گونه‌ای طراحی شده که تفسیر آن برای مراحل بعدی الگوریتم ساده‌تر باشد. امتیاز کیفیت هر توکن به عنوان منفی تابع تأثیر تعریف می‌شود:

$$Q(y_{i,j}|X_i, y_{i,:j}; \theta, \theta') = -\text{Inf}(y_{i,j}|X_i, y_{i,:j}; \theta, \theta') \quad (۳-۶)$$

این تعریف دارای منطق مشخص و هدفمندی است. با اعمال علامت منفی، امتیاز کیفیت مثبت نشان‌دهنده توکن‌هایی می‌شود که از آن‌ها به عنوان توکن‌های مفید و با کیفیت یاد می‌شود. این توکن‌ها کسانی هستند که مدل مرجع نسبت به مدل پایه قابلیت پیش‌بینی بهتری برای آن‌ها دارد، که نشان‌دهنده این است که آن‌ها الگوهای قابل تعمیم و یادگیری‌پذیری را نمایندگی می‌کنند.

توکن‌هایی که امتیاز کیفیت مثبت دارند، معمولاً شامل کلمات، عبارات، یا ساختارهای گرامری صحیح هستند که در زمینه مناسب خود قرار گرفته‌اند. این توکن‌ها ممکن است حاوی اطلاعات معنایی ارزشمند، الگوهای زبانی مفید، یا دانش تخصصی مرتبط با وظیفه مورد نظر باشند. یادگیری از این توکن‌ها انتظار می‌رود که به بهبود عملکرد کلی مدل و افزایش قابلیت‌های آن در انجام وظایف هدف کمک کند.

برعکس، امتیاز کیفیت منفی نشان‌دهنده توکن‌هایی است که ممکن است حاوی نویز، اطلاعات گمراه‌کننده، یا محتوای نامناسب باشند. این توکن‌ها ممکن است شامل خطاهای تایپی، کلمات اشتباه املا شده، ساختارهای گرامری نادرست، اطلاعات متناقض، یا حتی محتوایی باشند که در زمینه نامناسب قرار گرفته‌اند. حتی پس از آموزش اولیه، این توکن‌ها همچنان دشوار پیش‌بینی هستند یا عملکرد مدل در پیش‌بینی آن‌ها کاهش یافته است.

امتیازات کیفیت نزدیک به صفر نیز دسته خاصی از توکن‌ها را نمایندگی می‌کنند که نه کاملاً مفید و نه کاملاً مضر محسوب می‌شوند. این توکن‌ها ممکن است شامل کلمات رایج، ساختارهای زبانی ساده، یا محتوایی باشند که تأثیر قابل

توجهی بر عملکرد مدل ندارند. تصمیم‌گیری در مورد این دسته از توکن‌ها معمولاً بر اساس آستانه‌هایی که در مرحله تقسیم‌بندی تعیین می‌شوند، صورت می‌گیرد.

این سیستم امتیازدهی پایه محکمی برای مراحل بعدی الگوریتم فراهم می‌آورد و امکان تقسیم‌بندی هوشمند توکن‌ها به دسته‌های مثبت و منفی را فراهم می‌کند. با استفاده از این امتیازات، مدل قادر خواهد بود تصمیمات آگاهانه‌ای در مورد نحوه برخورد با هر توکن در فرآیند آموزش اتخاذ کند.

۳-۳ الگوریتم تقسیم‌بندی توکن‌ها

۳-۳-۱ اصول تقسیم‌بندی بر اساس امتیاز کیفیت

پس از محاسبه امتیازات کیفیت برای تمامی توکن‌های موجود در مجموعه داده آموزشی، گام بعدی تقسیم‌بندی این توکن‌ها به دو دسته متمایز مثبت و منفی است. این تقسیم‌بندی نقش محوری در عملکرد کلی سازوکار فراموشی ایفا می‌کند و باید بر اساس اصول آماری محکم و قابل اعتمادی صورت گیرد تا نتایج مطلوب حاصل شود.

فرآیند تقسیم‌بندی بر پایه این اصل بنیادی استوار است که توکن‌های دارای امتیاز کیفیت بالاتر نمایانگر اطلاعات مفیدتر و قابل اعتمادتری هستند که یادگیری از آن‌ها به بهبود عملکرد مدل کمک می‌کند. در مقابل، توکن‌های دارای امتیاز کیفیت پایین‌تر ممکن است حاوی نویز، اطلاعات گمراه‌کننده، یا محتوای نامناسبی باشند که نه تنها برای یادگیری مفید نیست، بلکه ممکن است تأثیر منفی بر فرآیند آموزش داشته باشد.

یکی از چالش‌های اساسی در این مرحله، تعیین معیار مناسب برای تقسیم‌بندی است. ساده‌ترین رویکرد ممکن استفاده از آستانه صفر است، به این صورت که توکن‌های دارای امتیاز مثبت در دسته مثبت و توکن‌های دارای امتیاز منفی در دسته منفی قرار گیرند. با این حال، تجربیات عملی نشان داده که این رویکرد همواره بهترین نتایج را ارائه نمی‌دهد، زیرا ممکن است منجر به عدم تعادل شدید بین دو دسته شود یا توکن‌هایی که امتیازشان نزدیک به صفر است، در دسته نامناسبی قرار گیرند.

بنابراین، در این پژوهش از رویکرد مبتنی بر صدک^۱ استفاده می‌شود که امکان کنترل بیشتری بر نسبت توکن‌های مثبت و منفی فراهم می‌آورد. این رویکرد بر اساس توزیع آماری امتیازات کیفیت در کل مجموعه داده عمل می‌کند و از پارامتر تنظیم‌پذیری به نام p استفاده می‌کند که درصد توکن‌هایی را که باید در دسته مثبت قرار گیرند، مشخص می‌کند.

^۱ Percentile

۳-۳-۲- تعیین آستانه تقسیم‌بندی

فرآیند تعیین آستانه تقسیم‌بندی شامل چندین مرحله دقیق و محاسباتی است که باید با توجه کامل به خصوصیات آماری داده‌ها انجام شود. ابتدا، تمامی امتیازات کیفیت محاسبه شده برای کل مجموعه داده در یک آرایه یا لیست جمع‌آوری می‌شوند. سپس، این امتیازات به ترتیب نزولی مرتب می‌شوند تا توزیع کاملی از بالاترین تا پایین‌ترین امتیاز به دست آید. با استفاده از پارامتر ρ که معمولاً مقداری بین $0/7$ تا $0/8$ انتخاب می‌شود، آستانه^۱ تقسیم‌بندی به صورت زیر محاسبه می‌گردد:

$$Threshold = FS(1 - \rho) \quad (۷-۳)$$

که در آن $FS(1 - \rho)$ نشان‌دهنده $(1 - \rho)$ امین صدک در توزیع مرتب‌شده امتیازات کیفیت است. به عبارت دیگر، اگر $\rho = 0/7$ باشد، آستانه برابر با ۳۰-امین صدک خواهد بود، که به این معناست که ۷۰ درصد از توکن‌ها دارای امتیاز بالاتر از این آستانه هستند و در دسته مثبت قرار می‌گیرند.

انتخاب مقدار مناسب برای ρ از اهمیت بالایی برخوردار است و تأثیر مستقیمی بر عملکرد نهایی الگوریتم دارد. مقادیر بالاتر ρ منجر به در نظرگیری درصد بیشتری از توکن‌ها به عنوان مثبت می‌شود، که ممکن است باعث شود برخی توکن‌های کم‌کیفیت نیز در دسته مثبت قرار گیرند. برعکس، مقادیر پایین‌تر ρ معیار سختگیرانه‌تری را اعمال می‌کند که ممکن است برخی توکن‌های مفید را در دسته منفی قرار دهد.

تجربیات عملی و آزمایش‌های گسترده نشان داده که بازه $0/7$ تا $0/8$ برای ρ معمولاً بهترین تعادل بین دقت تقسیم‌بندی و عملکرد نهایی را فراهم می‌آورد. با این حال، این مقدار ممکن است بسته به خصوصیات خاص مجموعه داده، نوع وظیفه، و مدل مورد استفاده نیاز به تنظیم داشته باشد.

۳-۳-۳ فرآیند طبقه‌بندی توکن‌ها

پس از تعیین آستانه، فرآیند طبقه‌بندی توکن‌ها به صورت ساده و مستقیم انجام می‌شود. برای هر توکن $y_{i,j}$ در مجموعه داده، امتیاز کیفیت آن با آستانه محاسبه شده مقایسه می‌گردد:

$$\mathcal{P} = (i, j) \in \mathcal{I} : Q(y_{i,j} | X_i, y_{i,:}; \theta, \theta') \geq FS(1 - \rho) \quad (۸-۳)$$

^۱ Threshold

$$\mathcal{N} = \mathcal{I} \setminus \mathcal{P}$$

که در آن \mathcal{P} مجموعه شاخص‌های توکن‌های مثبت و \mathcal{N} مجموعه شاخص‌های توکن‌های منفی را نمایش می‌دهد. نیز مجموعه کل شاخص‌های توکن‌ها در مجموعه داده است که قبلاً تعریف شده است. این تقسیم‌بندی باعث می‌شود که دقیقاً ρ درصد از توکن‌ها در دسته مثبت \mathcal{P} و $(1 - \rho)$ درصد باقی‌مانده در دسته منفی \mathcal{N} قرار گیرند. این توزیع کنترل‌شده امکان پیش‌بینی و مدیریت بهتر فرآیند آموزش را فراهم می‌آورد.

۳-۳-۴- ویژگی‌های الگوریتم تقسیم‌بندی

الگوریتم تقسیم‌بندی ارائه‌شده دارای چندین ویژگی مهم و مزیت قابل توجه است که آن را برای استفاده در مسائل عملی مناسب می‌سازد. نخست، این الگوریتم کاملاً قطعی و قابل تکرار است، به این معنا که با داده‌های ورودی یکسان همواره به نتایج مشابهی خواهد رسید. این ویژگی برای تضمین قابلیت اعتماد و مقایسه‌پذیری نتایج حائز اهمیت است. دوم، پیچیدگی محاسباتی این الگوریتم بسیار پایین است و عمدتاً شامل عملیات مرتب‌سازی و جستجو می‌شود که هر دو دارای پیچیدگی زمانی قابل قبولی هستند. حتی برای مجموعه داده‌های بزرگ، این فرآیند در زمان معقولی قابل اجرا است.

سوم، انعطاف‌پذیری بالای الگوریتم امکان تنظیم آسان آن برای انواع مختلف وظایف و مجموعه داده‌ها را فراهم می‌آورد. پارامتر ρ به عنوان تنها پارامتر تنظیم‌پذیر اصلی، کنترل ساده‌ای بر رفتار الگوریتم ارائه می‌دهد و برای تخمین این پارامتر، میتوان به نسبت تعداد توکن‌های مثبت نسبت به کل توکن‌ها توجه کرد که تخمین اولیه خیلی خوبی برای مقدار دهی اولیه این پارامتر است.

چهارم، الگوریتم نسبت به نویزهای جزئی در امتیازات کیفیت مقاوم است، زیرا بر اساس رتبه‌بندی کلی توکن‌ها عمل می‌کند نه مقادیر دقیق امتیازات. این ویژگی باعث می‌شود که خطاهای کوچک در محاسبه امتیازات تأثیر چندانی بر نتیجه نهایی نداشته باشند.

پنجم، این روش تضمین می‌کند که هیچ توکنی به طور کامل نادیده گرفته نمی‌شود، بلکه همه توکن‌ها در یکی از دو دسته مثبت یا منفی قرار می‌گیرند و در فرآیند آموزش شرکت می‌کنند، البته با نقش‌های متفاوت. این ویژگی یکی از مزایای اصلی رویکرد پیشنهادی نسبت به روش‌های سنتی انتخاب داده محسوب می‌شود.

۳-۳-۵- ملاحظات عملی و بهینه‌سازی

در پیاده‌سازی عملی الگوریتم تقسیم‌بندی، چندین ملاحظه مهم باید مورد توجه قرار گیرد. نخست، نحوه مدیریت حافظه هنگام کار با مجموعه داده‌های بزرگ اهمیت بالایی دارد. از آنجا که محاسبه و ذخیره امتیازات کیفیت برای میلیون‌ها توکن ممکن است حجم قابل توجهی از حافظه مصرف کند، استفاده از تکنیک‌های بهینه‌سازی حافظه مانند پردازش دسته‌ای یا ذخیره‌سازی فشرده ضروری است.

دوم، ترتیب پردازش داده‌ها می‌تواند تأثیری بر کارایی داشته باشد. معمولاً توصیه می‌شود که ابتدا تمامی امتیازات کیفیت محاسبه شده و سپس عملیات مرتب‌سازی و تقسیم‌بندی به صورت یکجا انجام شود تا از هماهنگی و سازگاری نتایج اطمینان حاصل شود.

سوم، امکان ذخیره و بارگذاری مجدد نتایج تقسیم‌بندی برای جلوگیری از تکرار محاسبات زمان‌بر مفید است. این ویژگی به ویژه در مراحل آزمایش و تنظیم پارامترها که ممکن است نیاز به اجرای مکرر الگوریتم باشد، اهمیت دارد. چهارم، اعتبارسنجی نتایج تقسیم‌بندی از طریق بررسی توزیع آماری دو دسته و مقایسه خصوصیات آن‌ها می‌تواند اطلاعات مفیدی در مورد کیفیت تقسیم‌بندی فراهم آورد. این بررسی شامل تحلیل میانگین، واریانس، و توزیع امتیازات در هر دسته می‌تواند باشد.

پنجم، مستندسازی دقیق پارامترها و تنظیمات استفاده شده در هر اجرا برای تضمین قابلیت تکرار و مقایسه نتایج ضروری است. این اطلاعات شامل مقدار ρ ، اندازه مجموعه داده مرجع، و جزئیات پیکربندی مدل‌های مورد استفاده باید ثبت شوند.

۳-۴- تابع هدف اصلاح‌شده و سازوکار فراموشی

پس از تکمیل فرآیند تقسیم‌بندی توکن‌ها به دو دسته مثبت و منفی، نیاز به طراحی تابع هدف نوینی احساس می‌شود که بتواند به‌طور همزمان از توکن‌های با کیفیت یادگیری کرده و توکن‌های گمراه‌کننده را به‌طور فعال فراموش کند. الگوریتم‌های تنظیم نظارت‌شده سنتی تمام توکن‌ها را یکسان در نظر گرفته و سعی در حداکثر کردن احتمال تولید همه آن‌ها دارند، که این رویکرد می‌تواند منجر به تقویت الگوهای نامطلوب و کاهش کیفیت کلی مدل شود. روش پیشنهادی با بهره‌گیری از سازوکار فراموشی انتخابی، رویکردی متفاوت و هدفمند را در پیش می‌گیرد که نه تنها از مقیاس کامل داده‌ها استفاده می‌کند، بلکه کیفیت یادگیری را نیز به‌طور قابل ملاحظه‌ای بهبود می‌بخشد.

۳-۴-۱- مبانی نظری فراموشی انتخابی

سازوکار فراموشی انتخابی بر پایه این اصل بنیادین استوار است که نه تمام اطلاعات موجود در مجموعه داده‌های آموزشی برای بهبود عملکرد مدل مفید هستند و نه نادیده‌گیری کامل اطلاعات نامطلوب بهترین راه حل محسوب می‌شود. این رویکرد از تکنیک‌های یادزدایی در یادگیری ماشین الهام گرفته و آن را برای زمینه تنظیم نظارت‌شده مدل‌های زبانی بزرگ تطبیق داده است. برخلاف روش‌های فیلتر کردن داده که اطلاعات نامطلوب را کاملاً حذف می‌کنند، فراموشی انتخابی از آن‌ها به عنوان سیگنال منفی برای راهنمایی فرآیند یادگیری استفاده می‌نماید.

اساس نظری این سازوکار بر این واقعیت استوار است که مدل‌های زبانی بزرگ دارای ظرفیت محدودی برای نگهداری و پردازش اطلاعات هستند. وقتی مدل مجبور به یادگیری الگوهای متضاد یا نامطلوب می‌شود، بخشی از این ظرفیت به اشتباه تخصیص یافته و از یادگیری الگوهای مفیدتر باز می‌ماند. فراموشی فعال با کاهش احتمال تولید توکن‌های نامطلوب، این ظرفیت را آزاد کرده و امکان تمرکز بیشتر بر الگوهای مطلوب را فراهم می‌آورد. این فرآیند منجر به تشکیل مرزهای دانشی واضح‌تری در مدل می‌شود که به بهبود تعمیم‌پذیری و کاهش بیش‌برازش کمک شایانی می‌کند.

علاوه بر این، فراموشی انتخابی به مدل کمک می‌کند تا درک بهتری از اینکه چه نوع پاسخ‌هایی نامطلوب هستند پیدا کند. این درک منفی به همان اندازه ارزشمند است که درک مثبت، زیرا به مدل امکان اجتناب فعال از تولید محتوای نامناسب را می‌دهد. در نتیجه، مدل نه تنها یاد می‌گیرد که چه چیزی را تولید کند، بلکه همچنین آموزش می‌بیند که از تولید چه محتوایی خودداری نماید.

۳-۴-۲- طراحی تابع هدف دوگانه

تابع هدف پیشنهادی بر مبنای ترکیب هوشمندانه دو هدف متمایز اما مکمل طراحی شده است. هدف اول، یادگیری سنتی توکن‌های مثبت از طریق حداکثر کردن احتمال تولید آن‌هاست، که مشابه تنظیم نظارت‌شده معمولی عمل می‌کند. هدف دوم، فراموش کردن فعال توکن‌های منفی از طریق حداقل کردن احتمال تولید آن‌هاست که نوآوری اصلی این پژوهش محسوب می‌شود. این دو هدف با استفاده از ضریب تعادل تطبیقی با یکدیگر ترکیب شده و تابع هدف نهایی را تشکیل می‌دهند.

مؤلفه یادگیری در تابع هدف مسئول پردازش توکن‌های مثبت است و عملکردی مشابه تنظیم نظارت‌شده استاندارد دارد. این مؤلفه میانگین وزن دار خطای^۱ محاسبه شده برای توکن‌های موجود در مجموعه مثبت را در نظر گرفته و سعی در

^۱ Error

کاهش آن دارد. مؤلفه فراموشی، که قلب نوآوری این روش محسوب می‌شود، میانگین وزن‌دار خطای محاسبه شده برای توکن‌های منفی را پردازش می‌کند، اما برخلاف مؤلفه یادگیری، این مقدار با علامت منفی در تابع هدف کل ظاهر می‌شود.

این طراحی دوگانه منجر به ایجاد یک فرآیند بهینه‌سازی چندهدفه می‌شود که پارامترهای مدل را همزمان در دو جهت هدایت می‌کند. گرادین‌های ناشی از مؤلفه یادگیری، مدل را به سمت بهبود تولید توکن‌های مثبت سوق می‌دهند، در حالی که گرادین‌های ناشی از مؤلفه فراموشی، مدل را از تولید توکن‌های منفی دور می‌نمایند. تعادل مناسب بین این دو، از طریق ضریب تطبیقی کنترل شده و امکان دستیابی به عملکرد بهینه را فراهم می‌آورد:

$$\mathcal{L}(\theta) = \frac{\sum_{(i,j) \in \mathcal{P}} w_{i,j} l(y_{i,j} | x_{i,j}; \theta)}{\sum_{(i,j) \in \mathcal{P}} w_{i,j}} - \lambda(\text{step}) \cdot \frac{\sum_{(i,j) \in \mathcal{N}} w_{i,j} l(y_{i,j} | x_{i,j}; \theta)}{\sum_{(i,j) \in \mathcal{N}} w_{i,j}} \quad (10-3)$$

در این فرمولاسیون، مؤلفه اول نشان‌دهنده میانگین خطای وزن‌دار بر روی توکن‌های مثبت و مؤلفه دوم نمایانگر میانگین خطای وزن‌دار بر روی توکن‌های منفی است که با ضریب $\lambda(\text{step})$ ضرب و از مؤلفه اول کسر می‌شود.

۳-۴-۳ سازوکار ضریب تعادل تطبیقی

یکی از چالش‌های کلیدی در پیاده‌سازی تابع هدف دوگانه، تعیین تعادل مناسب بین مؤلفه‌های یادگیری و فراموشی در طول فرآیند آموزش است. استفاده از ضریب ثابت ممکن است در مراحل مختلف آموزش عملکرد متفاوت و گاه نامطلوبی داشته باشد. مطالعات تجربی نشان داده‌اند که در مراحل ابتدایی آموزش، تمرکز بیشتر بر یادگیری توکن‌های مثبت مفیدتر است، در حالی که در مراحل انتهایی، تقویت سازوکار فراموشی برای تثبیت مرزهای دانشی و بهبود تعمیم‌پذیری اهمیت بیشتری پیدا می‌کند.

بر این اساس، ضریب تعادل $\lambda(\text{step})$ به صورت تطبیقی و بر اساس پیشرفت فرآیند آموزش تنظیم می‌شود. این ضریب در ابتدای آموزش مقدار پایینی داشته و به تدریج با پیشرفت آموزش افزایش می‌یابد. این رویکرد تطبیقی امکان بهره‌گیری بهینه از هر دو سازوکار یادگیری و فراموشی در زمان مناسب خود را فراهم می‌آورد. همچنین، این طراحی از ایجاد ناپایداری در مراحل ابتدایی آموزش جلوگیری کرده و امکان همگرایی تدریجی و پایدار را تضمین می‌کند. انتخاب تابع خطی برای تغییرات ضریب تعادل بر اساس مشاهدات تجربی و سادگی پیاده‌سازی صورت گرفته است. این

تابع امکان کنترل دقیق بر نرخ تغییرات و تنظیم آسان پارامترها را فراهم می‌آورد. علاوه بر این، رفتار خطی قابل پیش‌بینی بوده و امکان تحلیل و درک بهتر رفتار سیستم را فراهم می‌نماید.

$$\lambda(\text{step}) = t_{\min} + (t_{\max} - t_{\min}) \cdot \frac{\text{step}}{\text{total_steps}} \quad (11-3)$$

در این فرمول t_{\min} مقدار ابتدایی ضریب، t_{\max} مقدار نهایی آن، step مرحله فعلی آموزش و total_steps کل تعداد مراحل آموزش را نشان می‌دهد.

۳-۴-۴- پیاده‌سازی و جزئیات الگوریتم

پیاده‌سازی عملی سازوکار فراموشی انتخابی در قالب الگوریتم جامع و یکپارچه‌ای صورت می‌گیرد که چهار مرحله کلیدی را در برمی‌گیرد. مرحله اول شامل آموزش مدل مرجع بر روی زیرمجموعه‌ای از داده‌های آموزشی است که برای محاسبه امتیازات کیفیت توکن‌ها مورد استفاده قرار می‌گیرد. این مدل مرجع باید از کیفیت و پایداری مناسبی برخوردار باشد تا بتواند معیار قابل اعتمادی برای ارزیابی کیفیت توکن‌ها فراهم آورد.

مرحله دوم شامل ارزیابی جامع کیفیت تمام توکن‌های موجود در مجموعه داده آموزشی با استفاده از توابع تأثیر بین مدل پایه و مدل مرجع است. در این مرحله، برای هر توکن امتیاز کیفیت محاسبه شده و بر اساس این امتیازات، توکن‌ها به دو دسته مثبت و منفی تقسیم می‌شوند. مرحله سوم شامل انجام عملیات تقسیم‌بندی بر اساس پارامتر p است که تعیین می‌کند چه نسبتی از توکن‌ها در دسته مثبت قرار گیرند.

مرحله چهارم و نهایی، اجرای فرآیند آموزش با استفاده از تابع هدف دوگانه طراحی شده است. در هر مرحله از آموزش، ابتدا ضریب $\lambda(\text{step})$ بر اساس فرمول تطبیقی محاسبه می‌شود. سپس خطای میانگین برای توکن‌های مثبت و منفی به‌طور جداگانه محاسبه شده و تابع هدف کل تشکیل می‌شود. در نهایت، پارامترهای مدل با استفاده از الگوریتم بهینه‌سازی انتخابی به‌روزرسانی می‌شوند. این فرآیند تا رسیدن به همگرایی یا تکمیل تعداد مراحل آموزش تعریف شده ادامه می‌یابد.

فصل چهارم

نتایج و تحلیل یافته‌ها

در این فصل، نتایج حاصل از پیاده‌سازی و ارزیابی جامع روش فراموشی انتخابی برای بهبود تنظیم نظارت‌شده مدل‌های زبانی بزرگ ارائه می‌شود. برای تضمین اعتبار و قابلیت اطمینان نتایج، آزمایش‌های گسترده‌ای با استفاده از مجموعه داده‌های متنوع، مدل‌هایی با مقیاس‌های مختلف، و معیارهای ارزیابی جامع طراحی و اجرا شده است. این مطالعات شامل مقایسه عملکرد با روش‌های پایه، تحلیل حساسیت پارامترها، و بررسی تأثیر انتخاب‌های طراحی مختلف بر عملکرد نهایی سیستم می‌باشد.

۴-۱- تنظیمات آزمایشی و پیکربندی محیط

طراحی آزمایش‌های این پژوهش بر اساس اصول علمی دقیق و با رعایت استانداردهای بین‌المللی تحقیقات یادگیری ماشین صورت گرفته است. هدف اصلی این ارزیابی‌ها، تأیید اثربخشی روش پیشنهادی در شرایط متنوع و مقایسه آن با روش‌های موجود در ادبیات علمی است. برای تضمین اعتبار و قابلیت تکرارپذیری نتایج، تمام آزمایش‌ها با کنترل دقیق متغیرها و استفاده از دانه‌های ثابت و مختلف برای تولید اعداد تصادفی انجام شده‌اند.

۴-۱-۱- انتخاب و تنظیم مجموعه داده‌های آموزشی

مجموعه داده آموزشی این پژوهش از ترکیب هوشمندانه و متعادل پنج مجموعه داده معتبر و شناخته شده در حوزه تنظیم دستورالعمل‌ها تشکیل شده است. این انتخاب بر اساس تجربه محققان پیشین و نیاز به پوشش جامع انواع مختلف وظایف زبانی صورت گرفته است. هر یک از این مجموعه داده‌ها ویژگی‌ها و قوت‌های منحصربه‌فردی دارند که در مجموع، طیف وسیعی از چالش‌های موجود در تنظیم مدل‌های زبانی را پوشش می‌دهند.

ویژگی کلیدی	تعداد نمونه (10k)	تعداد نمونه (50k)	مجموعه داده
تنوع بالای وظایف	۳۵۹۳	۱۷۸۰۳	Flan_v2
دستورالعمل‌های پیچیده	۲۹۳۵	۱۴۳۴۴	WizardLM
کیفیت بالای حاشیه‌نویسی	۱۸۳۴	۹۲۷۶	Stanford Alpaca
رویکرد مکالمه‌ای	۱۱۳۵	۵۹۶۰	Open Assistant
داده‌های انسان‌ساخت	۵۰۳	۲۶۱۷	Dolly

جدول ۴-۱: توزیع دقیق مجموعه داده‌های آموزشی

Flan_v2 به عنوان ستون فقرات مجموعه داده آموزشی، شامل بیش از یک سوم کل نمونه‌هاست و طیف بسیار وسیعی از وظایف درک زبان طبیعی را پوشش می‌دهد. این مجموعه داده شامل وظایف مختلفی از جمله پرسش‌پاسخ، خلاصه‌سازی، ترجمه، طبقه‌بندی متن، و استدلال منطقی است که آن را به منبعی غنی و جامع برای آموزش مدل‌های چندمنظوره تبدیل می‌کند. WizardLM با تأکید ویژه بر دستورالعمل‌های پیچیده و چندمرحله‌ای، قابلیت مدل در پیروی از راهنمایی‌های پیچیده را تقویت می‌کند. این مجموعه داده به‌ویژه برای بهبود توانایی مدل در درک و اجرای وظایف پیچیده‌ای که نیاز به مراحل متوالی عملیات دارند، طراحی شده است.

Stanford Alpaca با تمرکز بر کیفیت بالای حاشیه‌نویسی و تنوع در ساختار دستورالعمل‌ها، الگوهای متنوعی از تعامل انسان-ماشین را ارائه می‌دهد. این مجموعه داده به گونه‌ای طراحی شده که مدل بتواند با انواع مختلف سبک‌های

ارتباطی و انتظارات کاربران سازگار شود. Open Assistant با رویکرد مکالمه‌ای و تأکید بر تعاملات چندنوبتی، به مدل کمک می‌کند تا قابلیت‌های گفتگوی طبیعی و پیوسته را توسعه دهد. این مجموعه داده شامل مکالمات پیچیده‌ای است که نیاز به حفظ زمینه و ارجاع به اطلاعات قبلی دارند.

Dolly با ارائه داده‌های تولید شده کاملاً توسط انسان، کیفیت و طبیعی بودن پاسخ‌ها را تضمین می‌کند. این مجموعه داده به عنوان معیاری برای کیفیت مطلوب در تولید متن عمل می‌کند و به مدل کمک می‌کند تا سبک نگارش طبیعی و قابل فهم انسان‌ها را فراگیرد. نسبت‌های انتخاب شده برای هر مجموعه داده بر اساس تحلیل تجربی و بررسی تأثیر هر یک بر عملکرد کلی مدل تعیین شده‌اند.

۴-۱-۲- معیارهای ارزیابی و مجموعه داده‌های آزمون

برای ارزیابی جامع و چندبعدی عملکرد روش پیشنهادی، پنج مجموعه داده معیار انتخاب شده که هر کدام جنبه خاصی از توانایی‌های مدل‌های زبانی بزرگ را مورد سنجش قرار می‌دهند. این انتخاب بر اساس پوشش جامع وظایف مختلف، تنوع در انواع سؤالات، و اعتبار علمی این معیارها در جامعه تحقیقاتی صورت گرفته است.

معیار ارزیابی	نوع پاسخ	طول سوال	اندازه داده	حوزه تمرکز	مجموعه داده
دقت	متن آزاد	متوسط	۸۱۷	صداقت و دقت اطلاعات	TruthfulQA
دقت	بله/خیر	کوتاه	۱۵۹۴۲	پرسش پاسخ بولین	BoolQ
دقت	چند گزینه‌ای	متوسط	۸۶۷۸	استدلال منطقی	LogiQA
F1-score	متن آزاد	متنوع	۲۰۴۰۰۰	پرسش پاسخ چندزبانه	TydiQA
دقت	عددی	متنوع	۲۳۰۵	حل مسائل کلامی ریاضی	ASDiv

جدول ۴-۲: مشخصات تفصیلی مجموعه داده‌های ارزیابی

TruthfulQA به عنوان معیاری برای سنجش توانایی مدل در ارائه اطلاعات صادقانه و دقیق طراحی شده است. این مجموعه داده شامل سؤالاتی است که اغلب مدل‌های زبانی در پاسخ به آن‌ها دچار خطاهای رایج یا تولید اطلاعات نادرست می‌شوند. ارزیابی بر اساس این معیار نشان‌دهنده قابلیت مدل در اجتناب از تولید محتوای گمراه‌کننده و حفظ صداقت در پاسخ‌هاست. BoolQ با ارائه سؤالات ساده با پاسخ بله/خیر، توانایی مدل در درک مفاهیم بنیادین و اتخاذ تصمیمات دودویی دقیق را مورد آزمون قرار می‌دهد.

LogiQA متمرکز بر ارزیابی قابلیت‌های استدلال منطقی مدل است و شامل مسائلی که نیاز به استنتاج منطقی، تحلیل روابط علت و معلول، و حل مسائل پیچیده دارند. این معیار به‌ویژه برای سنجش توانایی مدل در پردازش اطلاعات پیچیده و استخراج نتیجه‌گیری‌های منطقی اهمیت دارد. TydiQA با پوشش زبان‌های متنوع و متنوع بودن انواع سؤالات، قابلیت تعمیم‌پذیری مدل در زبان‌های مختلف را مورد ارزیابی قرار می‌دهد. این معیار نشان‌دهنده توانایی مدل در حفظ عملکرد مطلوب فراتر از زبان انگلیسی است.

ASDiv با تمرکز بر حل مسائل کلامی ریاضی، ترکیب دو مهارت مهم درک زبان طبیعی و محاسبات ریاضی را مورد آزمون قرار می‌دهد. این معیار به‌ویژه برای سنجش قابلیت مدل در تجزیه و تحلیل مسائل پیچیده و استخراج اطلاعات عددی از متن اهمیت دارد. ترکیب این پنج معیار پوشش جامعی از قابلیت‌های مورد انتظار از مدل‌های زبانی مدرن فراهم می‌آورد.

۴-۱-۳- انتخاب و پیکربندی مدل‌های آزمایشی

برای تضمین اعتبار و تعمیم‌پذیری نتایج، آزمایش‌ها بر روی چهار مدل زبانی بزرگ متن‌باز از خانواده LLaMA انجام شده که طیف وسیعی از اندازه‌ها و پیچیدگی‌ها را پوشش می‌دهند. انتخاب این مدل‌ها بر اساس دسترسی آزاد، عملکرد مطلوب در معیارهای استاندارد، و تنوع در اندازه‌ها صورت گرفته است تا امکان بررسی رفتار روش پیشنهادی در مقیاس‌های مختلف فراهم شود.

LLaMA-3.2-1B به‌عنوان کوچک‌ترین مدل در آزمایش‌ها، نماینده مدل‌هایی است که برای کاربردهای منابع محدود طراحی شده‌اند. این مدل امکان بررسی اثربخشی روش فراموشی در شرایطی که ظرفیت مدل محدود است را فراهم می‌آورد. LLaMA-3.2-3B با سه برابر پارامترهای مدل قبلی، نقطه تعادلی بین کارایی محاسباتی و عملکرد را نمایندگی می‌کند. این مدل معمولاً برای کاربردهای تجاری و صنعتی که نیاز به تعادل بین سرعت و کیفیت دارند، مورد استفاده قرار می‌گیرد.

LLaMA-3.1-8B به‌عنوان بزرگ‌ترین مدل در این مجموعه، نماینده مدل‌هایی است که برای کاربردهای تحقیقاتی و صنعتی پیشرفته طراحی شده‌اند. این مدل ظرفیت بالایی برای درک و پردازش اطلاعات پیچیده دارد و امکان بررسی حداکثر پتانسیل روش پیشنهادی را فراهم می‌آورد. علاوه بر این، مطالعات تکمیلی بر روی مدل LLaMA-2-13B نیز انجام شده تا قابلیت تعمیم‌پذیری روش در مقیاس‌های بزرگ‌تر تأیید شود.

تهیه مدل‌های مرجع با دقت ویژه‌ای انجام شده است. این مدل‌ها از طریق تنظیم نظارت‌شده مدل‌های پایه بر روی زیرمجموعه‌ای از داده‌های آموزشی تهیه شده‌اند که هیچ تداخلی با مجموعه داده آموزشی اصلی ندارند. این جداسازی

کامل تضمین می‌کند که محاسبه امتیازات تأثیر بر اساس اطلاعات کاملاً مستقل صورت گرفته و از هرگونه تورش در ارزیابی کیفیت توکن‌ها جلوگیری شود. همچنین، کیفیت مدل‌های مرجع از طریق ارزیابی بر روی مجموعه داده‌های معیار تأیید شده است.

۴-۱-۴- تنظیمات فنی و پیکربندی آموزش

پیاده‌سازی تنظیم نظارت‌شده با استفاده از تکنیک LoRA انجام شده که انتخابی استراتژیک برای تعادل بین کارایی محاسباتی و کیفیت آموزش محسوب می‌شود. پارامترهای LoRA با دقت تنظیم شده‌اند: اندازه رتبه^۱ برابر ۶۴ که تعادل مناسبی بین قدرت نمایندگی و کارایی فراهم می‌آورد، ضریب مقیاس‌بندی^۲ برابر ۱۶ که میزان تأثیر تطبیق‌های کم‌رنگ را کنترل می‌کند، و نرخ حذف تصادفی^۳ برابر ۰/۰۱ که از بیش‌برازش جلوگیری می‌کند.

انتخاب بهینه‌ساز AdamW بر اساس عملکرد مطلوب آن در آموزش مدل‌های زبانی بزرگ صورت گرفته است. اندازه کل دسته^۴ برابر ۲۴ به گونه‌ای انتخاب شده که حداکثر استفاده از منابع محاسباتی موجود را تضمین کند و در عین حال ثبات آموزش را حفظ نماید. نرخ یادگیری ۰/۰۰۰۱ بر اساس آزمایش‌های اولیه و بررسی ادبیات مربوطه تعیین شده است. برنامه‌ریز نرخ یادگیری خطی با نسبت warm-up برابر ۰/۰۳ برای تضمین شروع نرم و تدریجی فرآیند آموزش استفاده شده است.

محیط سخت‌افزاری با در نظر گیری نیازهای محاسباتی سنگین مدل‌های زبانی بزرگ طراحی شده است. چهار GPU از نوع NVIDIA L40S-48GB که هر کدام دارای ۴۸ گیگابایت حافظه اختصاصی هستند، قدرت پردازش موازی لازم برای آموزش کارآمد مدل‌ها را فراهم می‌آورند. پردازنده‌های Intel Xeon 6338 با معماری پیشرفته، پردازش CPU مورد نیاز برای عملیات‌های تکمیلی و مدیریت داده‌ها را تأمین می‌کنند.

سیستم‌عامل Ubuntu 20.04.6 LTS به‌عنوان پلتفرم پایه انتخاب شده که ثبات و سازگاری بالایی با ابزارهای یادگیری ماشین دارد. کتابخانه Transformers نسخه ۴.۵۱.۳ به‌عنوان چارچوب اصلی برای پیاده‌سازی مدل‌ها و CUDA نسخه ۱۲.۵ برای بهره‌گیری بهینه از قدرت پردازش GPU ها استفاده شده‌اند. این ترکیب نرم‌افزاری پایداری، کارایی و سازگاری مطلوبی را برای انجام آزمایش‌های پیچیده فراهم می‌آورد.

زمان‌های آموزش اندازه‌گیری شده نشان‌دهنده قابلیت اجرای عملی روش پیشنهادی هستند. مدل ۱ میلیارد پارامتری در

^۱ rank

^۲ Scaling factor

^۳ dropout

^۴ batch

حدود ۲ ساعت، مدل ۳ میلیارد پارامتری در حدود ۳ ساعت، و مدل ۸ میلیارد پارامتری در حدود ۵ ساعت آموزش می‌بینند. این زمان‌ها نشان می‌دهند که روش پیشنهادی بدون افزایش قابل توجه در زمان آموزش نسبت به تنظیم نظارت‌شده استاندارد قابل اجرا است.

ارزیابی‌های نهایی با استفاده از کتابخانه معتبر lm-evaluation-harness انجام شده که استاندارد شناسایی شده برای ارزیابی مدل‌های زبانی محسوب می‌شود. این کتابخانه تضمین می‌کند که ارزیابی‌ها با روش‌های استاندارد و قابل مقایسه انجام شوند. همه آزمایش‌ها سه بار با دانه‌های مختلف تکرار شده و میانگین همراه با انحراف معیار گزارش شده تا پایداری و قابلیت اطمینان نتایج نشان داده شود.

۴-۲- نتایج مقایسه‌ای و تحلیل عملکرد

این بخش نتایج جامع حاصل از ارزیابی روش فراموشی انتخابی پیشنهادی در مقایسه با روش‌های پایه و رقیب را ارائه می‌دهد. همه آزمایش‌ها با سه دانه تصادفی مختلف تکرار شده و میانگین نتایج همراه با انحراف معیار گزارش شده است تا پایداری و قابلیت اطمینان روش پیشنهادی نشان داده شود. نتایج نشان‌دهنده بهبود قابل توجه و مداوم عملکرد در تمام مدل‌ها و معیارهای ارزیابی است.

۴-۲-۱ مقایسه عملکرد بر روی مدل‌های مختلف اندازه

جدول زیر نتایج مقایسه‌ای روش پیشنهادی با چهار روش پایه بر روی سه مدل LLaMA با اندازه‌های مختلف را نشان می‌دهد. روش‌های مقایسه شامل: مدل پایه بدون تغییر، تنظیم نظارت‌شده استاندارد روی تمام توکن‌ها، نادیده‌گیری توکن‌های منفی (نادیده‌گیری^۱)، و روش فراموشی پیشنهادی می‌باشند.

¹ Ignoring

روش	TruthfulQA	BoolQ	LogiQA	TydiQA	ASDiv	میانگین
LLaMA-3.2-1B						
پایه	۳۷.۸۳±۰	۶۳.۸۰±۰	۲۲.۱۷±۰	۱۴.۳۶±۰	۰±۰	۲۷.۶۳±۰
تمام ترکی	۳۸.۷۴±۰.۳۹	۵۹.۸۴±۰.۹۴	۲۴.۶۰±۰.۲۵	۲۸.۱۰±۰.۴۶	۰.۵۵±۰.۴۸	۳۰.۳۷±۰.۳۹
نادیده گیری	۴۲.۴۰±۰.۱۳	۶۰.۲۱±۱.۶۶	۲۴.۳۴±۰.۳۱	۳۳.۸۷±۰.۶۴	۰.۹۱±۰.۲۰	۳۲.۳۵±۰.۴۶
فراموشی	۴۴.۸۳±۰.۴۵	۶۵.۳۹±۰.۳۹	۲۵.۶۰±۰.۴۸	۳۶.۲۱±۰.۷۷	۲.۲۸±۰.۰۴	۳۴.۸۶±۰.۲۲
LLaMA-3.2-3B						
پایه	۳۹.۴۵±۰	۷۳.۰۴±۰	۲۲.۱۷±۰	۲۱.۱۲±۰	۳۱.۲۴±۰	۳۷.۴۰±۰
تمام ترکی	۴۲.۹۵±۰.۴۷	۷۲.۵۴±۰.۵۹	۲۵.۵۱±۰.۲۱	۴۴.۰۴±۰.۲۷	۴۹.۴۶±۰.۱۴	۴۶.۹۰±۰.۱۶
نادیده گیری	۴۷.۲۳±۰.۸۶	۷۵.۴۰±۰.۳۷	۲۵.۱۲±۰.۳۱	۴۷.۶۳±۰.۴۲	۴۸.۵۱±۰.۷۴	۴۸.۷۸±۰.۱۹
فراموشی	۵۰.۳۲±۰.۹۶	۷۶.۶۶±۰.۰۷	۲۷.۰۹±۰.۳۷	۵۶.۳۶±۰.۰۶	۵۰.۴۷±۰.۳۰	۵۲.۱۸±۰.۱۲
LLaMA-3.1-8B						
پایه	۴۵.۰۸±۰	۸۲.۱۵±۰	۲۶.۵۱±۰	۴۶.۶۷±۰	۱۲.۹۳±۰	۴۲.۶۷±۰
تمام ترکی	۴۴.۵۱±۰.۴۸	۸۱.۴۴±۰.۴۷	۲۵.۶۸±۰.۱۴	۵۲.۰۳±۰.۱۸	۵۱.۴۶±۰.۴۲	۵۱.۰۲±۰.۱۱
نادیده گیری	۵۲.۳۸±۰.۲۲	۸۲.۷۶±۰.۰۷	۲۵.۵۳±۰.۱۱	۵۶.۶۶±۰.۰۶	۵۷.۹۵±۰.۳۵	۵۵.۰۶±۰.۱۶
فراموشی	۵۸.۳۹±۰.۶۵	۸۳.۱۴±۰.۱۵	۳۱.۱۵±۰.۸۶	۶۶.۲۱±۰.۲۳	۵۷.۴۸±۰.۱۲	۵۹.۲۷±۰.۳۵

جدول ۳-۴: مقایسه عملکرد روش های مختلف بر روی مدل های LLaMA

نتایج نشان می دهند که روش فراموشی در تمام مدل ها و تقریباً تمام معیارها عملکرد برتری دارد. برای مدل-LLaMA 3.2-1B، روش پیشنهادی بهبود ۴/۴۹٪ نسبت به تنظیم نظارت شده استاندارد و ۲/۵۱٪ نسبت به نادیده گیری در سطح توکن حاصل کرده است. این بهبودها در مدل های بزرگ تر قابل توجه تر می شوند، به طوری که در مدل-LLaMA 3.1-8B بهبود ۸/۲۵٪ نسبت به تنظیم نظارت شده استاندارد مشاهده می شود. همچنین برای تأیید قابلیت تعمیم پذیری روش پیشنهادی در مقیاس های بزرگ تر، آزمایش های تکمیلی بر روی مدل LLaMA-2-13B انجام شده است. نتایج زیر نشان دهنده حفظ اثربخشی روش در مدل های بزرگ تر است.

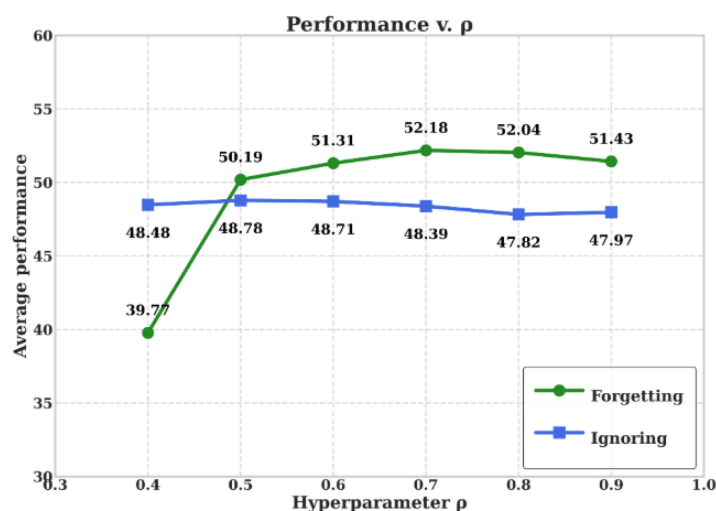
میانگین	ASDiv	TydiQA	LogiQA	BoolQ	TruthfulQA	روش
۳۵/۶۱	۰/۳۵	۳۴/۲۷	۲۶/۰۵	۸۰/۶۷	۳۶/۷۳	پایه
۳۹/۵۷	۸/۷۶	۳۶/۷۷	۲۷/۴۴	۸۲/۲۴	۴۲/۶۵	تمام ترکی
۴۱/۷۱	۱۵/۳۴	۳۸/۳۹	۲۷/۲۹	۸۴/۵۰	۴۳/۰۱	نادیده گیری
۴۶/۲۸	۱۷/۸۰	۴۸/۷۱	۲۷/۹۵	۸۴/۱۳	۵۲/۸۲	فراموشی

جدول ۴-۲: مقایسه عملکرد بر روی مدل LLaMA-2-13B

نتایج مدل ۱۳ میلیارد پارامتری تأیید می کند که روش فراموشی حتی در مقیاس های بزرگ تر نیز اثربخش است.

۴-۲-۲- تحلیل حساسیت پارامترها

یکی از جنبه های حیاتی ارزیابی هر روش پیشنهادی، بررسی حساسیت آن نسبت به پارامترهای کلیدی است. در این بخش، تأثیر پارامتر ρ که نسبت تقسیم بندی توکن های مثبت و منفی را تعیین می کند، مورد بررسی دقیق قرار می گیرد تا درک بهتری از رفتار روش پیشنهادی در شرایط مختلف حاصل شود.



شکل ۴-۱: تحلیل عملکرد روش فراموشی در مقابل نادیده گیری در مقادیر مختلف ρ

شکل ۴-۱ مقایسه جامع عملکرد میانگین روش فراموشی پیشنهادی و روش نادیده گیری در طیف وسیعی از مقادیر ρ را

نشان می‌دهد. نتایج حاصل از این تحلیل یافته‌های قابل توجهی را آشکار می‌کند که نه تنها برتری روش پیشنهادی را تأیید می‌کند، بلکه ثبات و مقاومت آن را در برابر تغییرات پارامترها نیز نشان می‌دهد. همان‌طور که در نمودار مشاهده می‌شود، روش فراموشی در تمام مقادیر آزمایش شده p عملکرد قابل توجه و مداومی برتر از روش نادیده‌گیری ارائه می‌دهد. این برتری در نقطه $p = 0.7$ به اوج خود می‌رسد، جایی که روش فراموشی به عملکرد $52/18\%$ دست می‌یابد در حالی که روش نادیده‌گیری تنها $48/39\%$ کسب می‌کند، که نشان‌دهنده اختلاف قابل توجه $3/79\%$ بین دو روش است. علاوه بر این، هر دو روش در بازه 0.6 تا 0.8 عملکرد نسبتاً پایداری نشان می‌دهند که اهمیت ویژه‌ای دارد، زیرا این ثبات نشان می‌دهد که روش پیشنهادی نسبت به تغییرات جزئی در انتخاب پارامتر مقاوم است و امکان استفاده عملی آن را در شرایط مختلف تضمین می‌کند و حتی با استفاده از نسبت تعداد کل توکن‌های مثبت به کل تعداد توکن‌ها مقدار دهی اولیه مناسبی برای این پارامتر داشت. در مقابل، در مقادیر پایین p ، به‌ویژه در $p = 0.4$ ، هر دو روش کاهش قابل ملاحظه‌ای در عملکرد تجربه می‌کنند که این پدیده به دلیل قرارگیری تعداد زیادی از توکن‌های مفید و ارزشمند در دسته منفی است، که منجر به فراموشی یا نادیده‌گیری اطلاعات حیاتی می‌شود و در نتیجه قابلیت‌های کلی مدل را تضعیف می‌کند. نکته قابل توجه دیگر این است که اختلاف عملکرد بین دو روش در سراسر طیف مقادیر p تقریباً ثابت باقی می‌ماند، که این امر نشان‌دهنده برتری بنیادین و اساسی سازوکار فراموشی فعال نسبت به نادیده‌گیری ساده است و تأیید می‌کند که مزیت روش پیشنهادی نه تنها در یک نقطه خاص، بلکه در طیف وسیعی از تنظیمات پارامتر قابل مشاهده است.

پارامترهای مهم دیگر در این پژوهش، t_{min} و t_{max} هستند که نتایج آن در جدول ۵-۴ ارائه شده است. این بررسی با ثابت نگه داشتن پارامتر p در مقدار 0.7 و آزمایش ترکیبات مختلف از مقادیر t_{min} و t_{max} بر روی مدل LLaMA-3.2-3B صورت گرفته است. نتایج حاصل از این تحلیل یافته‌های قابل توجهی را آشکار می‌کند که نشان‌دهنده ثبات و مقاومت قابل ملاحظه روش پیشنهادی است. همان‌طور که از جدول مشخص است، تغییرات عملکرد در سراسر ترکیبات مختلف پارامترها در محدوده نسبتاً محدودی قرار دارد، به طوری که انحراف معیار کل نتایج کمتر از 0.5% است که این امر نشان‌دهنده پایداری بالای روش پیشنهادی است. بهترین عملکرد در ترکیب $t_{min} = 0.0001$ و $t_{max} = 0.25$ با میانگین $52/18\%$ حاصل شده که همان تنظیماتی است که در آزمایش‌های اصلی استفاده شده است، اما قابل توجه است که سایر ترکیبات نیز عملکرد نزدیکی به این مقدار بهینه ارائه می‌دهند. بررسی دقیق‌تر نتایج نشان می‌دهد که پارامتر t_{min} در بازه 0.0001 تا 0.01 و پارامتر t_{max} در بازه 0.15 تا 0.45 تغییرات قابل ملاحظه‌ای در عملکرد کلی ایجاد نمی‌کنند، که این امر اهمیت ویژه‌ای برای کاربردهای عملی دارد زیرا نشان می‌دهد که تنظیم دقیق این پارامترها الزامی نیست و روش پیشنهادی در طیف وسیعی از تنظیمات عملکرد مطلوبی ارائه می‌دهد. همچنین، این ثبات تا حدودی به دلیل مقاومت ذاتی مدل‌های زبانی بزرگ و دانش گسترده پیش‌آموزشی آن‌هاست که پایه محکمی برای

مقاومت در برابر تغییرات متوسط در پارامترهای تنظیم نظارت شده فراهم می‌آورد. این یافته‌ها اطمینان می‌دهند که روش پیشنهادی نیاز به تنظیم دقیق و پیچیده پارامترها ندارد و می‌تواند با تنظیمات نسبتاً ساده و استاندارد به عملکرد مطلوب دست یابد.

تmin	تmax	TruthfulQA	BoolQ	LogiQA	TydiQA	ASDiv	میانگین
۰/۱۰۰۰۰	۰/۵۴	۵۲/۷۵	۷۴/۳۸	۲۵/۸۹	۵۴/۲۷	۴۸/۱۰	۵۱/۰۸
۰/۱۰۰۰۰	۰/۵۳	۵۱/۵۵	۷۵/۱۱	۲۶/۱۵	۵۶/۷۴	۴۸/۴۲	۵۱/۵۹
۰/۱۰۰۰۰	۰/۵۲	۵۰/۹۳	۷۶/۵۸	۲۵/۹۹	۵۶/۱۳	۵۰/۲۶	۵۱/۹۸
۰/۱۰۰۰۰	۰/۵۱	۵۰/۱۷	۷۵/۴۵	۲۶/۱۹	۵۴/۳۷	۵۰/۶۷	۵۱/۳۷
۰/۱۰۰۰	۰/۵۴	۵۰/۹۰	۷۷/۵۶	۲۵/۸۳	۵۴/۳۳	۴۸/۹۰	۵۱/۵۰
۰/۱۰۰۰	۰/۵۳	۵۱/۲۰	۷۵/۶۷	۲۶/۶۵	۵۷/۲۱	۴۸/۷۸	۵۱/۹۰
۰/۱۰۰۰	۰/۵۲	۵۰/۳۲	۷۶/۶۴	۲۷/۰۹	۵۶/۳۶	۵۰/۴۷	۵۲/۱۸
۰/۱۰۰۰	۰/۵۱	۵۰/۰۹	۷۴/۷۹	۲۵/۲۷	۵۵/۲۱	۵۱/۸۲	۵۱/۴۴
۰/۱۰۰	۰/۵۱	۴۹/۰۵	۷۶/۰۳	۲۶/۳۶	۵۴/۸۵	۵۱/۴۹	۵۱/۵۶
۰/۱۰۰	۰/۵۲	۴۸/۹۶	۷۶/۵۰	۲۸/۶۸	۵۶/۳۵	۴۹/۶۶	۵۲/۰۳
۰/۱۰۰	۰/۵۳	۵۱/۲۵	۷۴/۴۱	۲۶/۵۱	۵۶/۵۸	۵۰/۰۵	۵۱/۷۶
۰/۱۰۰	۰/۵۴	۵۰/۶۹	۷۴/۵۰	۲۵/۹۸	۵۶/۹۷	۴۸/۴۶	۵۱/۳۲
۰/۱۰	۰/۵۱	۵۰/۴۶	۷۵/۲۴	۲۶/۱۲	۵۴/۱۷	۵۰/۹۵	۵۱/۳۹
۰/۱۰	۰/۵۲	۵۱/۰۲	۷۶/۲۸	۲۷/۷۵	۵۵/۴۸	۴۹/۹۳	۵۲/۰۹
۰/۱۰	۰/۵۳	۵۲/۷۸	۷۴/۴۴	۲۵/۵۸	۵۵/۹۲	۴۸/۳۰	۵۱/۴۰
۰/۱۰	۰/۵۴	۵۰/۰۹	۷۴/۸۷	۲۷/۶۰	۵۴/۶۹	۴۸/۶۸	۵۱/۱۹

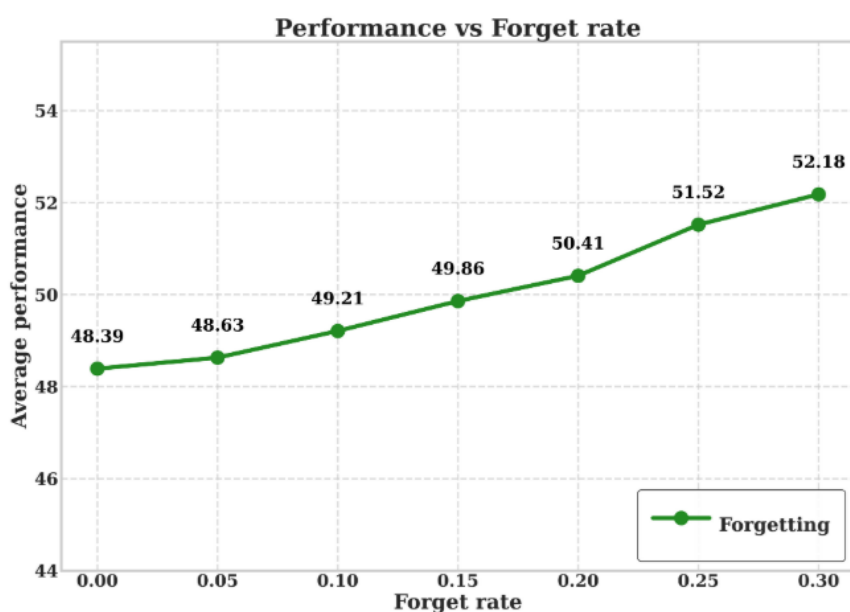
جدول ۴-۵: تحلیل حساسیت پارامترهای tmin و tmax

۴-۲-۳- تأثیر نرخ فراموشی بر عملکرد

علاوه بر بررسی تأثیر پارامتر p ، تحلیل رفتار روش پیشنهادی در مقابل نرخ‌های مختلف فراموشی^۱ نیز اهمیت بسزایی دارد. شکل ۲-۴ نمودار عملکرد میانگین روش فراموشی پیشنهادی را در برابر طیف وسیعی از نرخ‌های فراموشی نشان

^۱ Forget rate

می‌دهد که یافته‌های قابل توجه و بصیرت‌انگیزی را در مورد رفتار این سازوکار آشکار می‌کند. همان‌طور که در نمودار مشاهده می‌شود، عملکرد روش پیشنهادی رابطه مستقیم و پیوسته‌ای با نرخ فراموشی نشان می‌دهد، به طوری که با افزایش تدریجی نرخ فراموشی از ۰/۰ به ۰/۳، عملکرد به طور مداوم و قابل ملاحظه‌ای بهبود می‌یابد. در نقطه شروع، یعنی زمانی که نرخ فراموشی برابر ۰/۰ است که عملاً معادل عدم استفاده از سازوکار فراموشی محسوب می‌شود، عملکرد ۴۸/۳۹٪ حاصل می‌شود، اما با افزایش تدریجی این نرخ، شاهد بهبود پیوسته و قابل توجهی هستیم که در نرخ فراموشی ۰/۳۰ به اوج خود یعنی ۵۲/۱۸٪ می‌رسد. این روند صعودی نشان می‌دهد که فراموشی فعال توکن‌های منفی نه تنها تأثیر مثبت بر عملکرد دارد، بلکه این تأثیر با فرض ثابت بودن p ، با افزایش شدت فراموشی تقویت می‌شود، که این یافته با فرضیه نظری روش پیشنهادی مبنی بر اینکه حذف فعال توکن‌های گمراه‌کننده موجب آزادسازی ظرفیت مدل برای یادگیری الگوهای مفیدتر می‌شود، کاملاً همخوانی دارد. همچنین، شکل تدریجی و هموار منحنی نشان می‌دهد که روش پیشنهادی رفتار پیش‌بینی‌پذیر و کنترل‌شده‌ای دارد و احتمال وجود نقاط بحرانی یا تغییرات ناگهانی در عملکرد که ممکن است منجر به ناپایداری شود، بسیار پایین است.



شکل ۴-۲: تحلیل عملکرد میانگین روش فراموشی با نرخ‌های مختلف فراموشی

۴-۲-۴- تحلیل الگوهای بهبود

بررسی دقیق‌تر نتایج الگوهای جالبی را آشکار می‌کند. بهبودهای حاصل از روش پیشنهادی با افزایش اندازه مدل تشدید

می‌شوند، که نشان‌دهنده مقیاس‌پذیری مطلوب این روش است. در مدل ۱ میلیارد پارامتری بهبود متوسط ۴/۴۹٪ مشاهده می‌شود، در حالی که در مدل ۸ میلیارد پارامتری این بهبود به ۸/۲۵٪ می‌رسد.

معیار TruthfulQA به‌طور مداوم بیشترین بهبود را نشان می‌دهد، که ممکن است نشان‌دهنده اثربخشی ویژه روش پیشنهادی در کاهش تولید اطلاعات نادرست باشد. این یافته با هدف اصلی سازوکار فراموشی که حذف توکن‌های گمراه‌کننده است، همخوانی دارد. در مقابل، معیارهایی مانند LogiQA که نیاز به استدلال پیچیده دارند، بهبودهای متعادل‌تری نشان می‌دهند.

۴-۲-۵- پایداری و تکرارپذیری نتایج

انحراف معیار پایین در تمام آزمایش‌ها نشان‌دهنده پایداری بالای روش پیشنهادی است. در اکثر موارد، انحراف معیار کمتر از ۱٪ است که نشان می‌دهد نتایج قابل اعتماد و تکرارپذیر هستند. این پایداری برای کاربردهای عملی و صنعتی روش پیشنهادی اهمیت بالایی دارد.

همچنین روش پیشنهادی بدون افزایش قابل توجه در هزینه محاسباتی نسبت به تنظیم نظارت‌شده استاندارد قابل اجرا است. این ویژگی روش پیشنهادی را برای استفاده عملی جذاب‌تر می‌کند، زیرا بهبود عملکرد بدون افزایش قابل توجه منابع مورد نیاز حاصل می‌شود.

فصل پنجم

نتیجه‌گیری و پیشنهادات برای تحقیقات آتی

در این فصل، تحلیل جامع و نتیجه‌گیری نهایی از پژوهش انجام‌شده در حوزه فراموشی انتخابی برای بهبود تنظیم نظارت‌شده مدل‌های زبانی بزرگ ارائه می‌شود. پس از بررسی مبانی نظری، طراحی روش‌شناسی، و ارزیابی تجربی در فصول پیشین، این فصل به استخراج دستاوردهای کلیدی، شناسایی محدودیت‌های موجود، و تعیین مسیرهای آتی تحقیق اختصاص دارد. هدف اصلی، ارائه چشم‌اندازی جامع از جایگاه روش پیشنهادی در ادبیات علمی و تأثیرات آن بر پیشرفت حوزه مدل‌های زبانی بزرگ است. این تحلیل شامل جمع‌بندی یافته‌های تجربی، بررسی چالش‌های باقی‌مانده، و ارائه پیشنهادات مشخص برای توسعه و بهبود بیشتر سازوکار فراموشی انتخابی می‌باشد.

۵-۱- نتیجه‌گیری و دستاوردهای کلیدی

این پژوهش سازوکار "فراموشی انتخابی" را برای بهبود تنظیم نظارت‌شده مدل‌های زبانی بزرگ ارائه داده است. هدف اصلی حل مسئله وابستگی شدید این مدل‌ها به کیفیت داده‌ها و کاهش تأثیر منفی توکن‌های گمراه‌کننده بوده است. برخلاف روش‌های سنتی که داده‌های نامطلوب را حذف می‌کنند، روش پیشنهادی از آن‌ها به‌عنوان سیگنال منفی برای بهبود یادگیری استفاده می‌کند.

دستاوردهای کلیدی شامل طراحی چارچوب تقسیم‌بندی توکن‌ها بر اساس توابع تأثیر، ارائه تابع هدف دوگانه با ضریب تعادل تطبیقی، و اثبات برتری رویکرد سطح توکن نسبت به سطح جمله است. نتایج تجربی بر روی مدل‌های LLaMA نشان می‌دهد که روش پیشنهادی بهبودهای ۴۹٪ تا ۸۲۵٪ نسبت به تنظیم نظارت‌شده استاندارد حاصل می‌کند. این بهبودها با افزایش اندازه مدل تشدید می‌شوند که نشان‌دهنده مقیاس‌پذیری مطلوب است.

نوآوری اصلی در تغییر نگاه به داده‌های نامطلوب از "مانعی برای حذف" به "منبعی برای بهبود یادگیری" است. این رویکرد امکان حفظ مقیاس کامل داده‌ها را فراهم می‌آورد و مرزهای دانشی واضح‌تری در مدل ایجاد می‌کند. تحلیل حساسیت پارامترها نشان می‌دهد که روش در طیف وسیعی از تنظیمات عملکرد پایداری دارد که برای کاربردهای عملی اهمیت دارد.

۵-۲- محدودیت‌ها و چالش‌های موجود

روش پیشنهادی محدودیت‌هایی دارد که باید در نظر گرفته شوند. اولین محدودیت وابستگی به کیفیت مدل مرجع است، زیرا ارزیابی توکن‌ها بر اساس مقایسه با این مدل صورت می‌گیرد. دومین محدودیت، حساسیت به اندازه مجموعه داده و نسبت نویز است که در داده‌های کوچک یا کم‌نویز اثربخشی را کاهش می‌دهد. سومین محدودیت، پیچیدگی اضافی در فرآیند آموزش و نیاز به تنظیم پارامترهای بیشتری است. چهارمین محدودیت، عدم اطمینان از رفتار روش در حوزه‌های تخصصی خاص مانند پزشکی یا حقوق که نیاز به دقت بالا دارند. همچنین، انتخاب نامناسب پارامتر p می‌تواند منجر به فراموشی اطلاعات مفید یا حفظ توکن‌های نامطلوب شود.

۵-۳- پیشنهادات برای تحقیقات آتی

بر اساس یافته‌ها، چندین مسیر برای توسعه آتی پیشنهاد می‌شود. اول، توسعه روش‌های خودکار برای تنظیم پارامترهای کلیدی که بتوانند بهترین مقادیر را بر اساس ویژگی‌های مجموعه داده تعیین کنند. دوم، بررسی اثربخشی روش در حوزه‌های تخصصی مختلف برای درک بهتر قابلیت‌ها و محدودیت‌های آن در کاربردهای واقعی. سوم، توسعه معیارهای پیشرفته‌تر برای ارزیابی کیفیت توکن‌ها با استفاده از تکنیک‌های یادگیری عمیق جدیدتر یا ترکیب چندین معیار. چهارم، گسترش روش به سایر مراحل چرخه حیات مدل‌های زبانی شامل پیش‌آموزش و بهینه‌سازی تریجیات. پنجم، ترکیب روش فراموشی با سایر تکنیک‌های بهبود داده‌ها مانند تولید داده مصنوعی یا تقویت داده‌ها.

ششم، مطالعه تأثیرات بلندمدت فراموشی بر قابلیت‌های مدل و تعیین استراتژی‌هایی برای حفظ تعادل بین فراموشی و حفظ دانش پیشین. هفتم، توسعه ابزارهای عملی و متن‌باز برای پیاده‌سازی آسان روش که امکان استفاده گسترده توسط محققان و توسعه‌دهندگان را فراهم آورد.

در نهایت، این پژوهش نشان می‌دهد که با نگاه نوآورانه به داده‌های نامطلوب، می‌توان بهبودهای قابل توجهی در عملکرد مدل‌های زبانی حاصل کرد. سازوکار فراموشی انتخابی چشم‌اندازی جدید برای استفاده هوشمندانه از تمام داده‌های موجود ارائه می‌دهد و مسیری برای تحقیقات آتی در جهت توسعه روش‌های کارآمدتر آموزش مدل‌های زبانی فراهم می‌کند.

مراجع

- [1] Brown, T., Mann, B., Ryder, N., et al. Language models are few-shot learners. NeurIPS, 2020.
- [2] Dodge, J., Sap, M., Marasovic, A., et al. Documenting large webtext corpora: A case study on the colossal clean crawled corpus. EMNLP, 2021.
- [3] Zhou, C., Liu, P., Xu, P., et al. LIMA: Less is more for alignment. NeurIPS, 2023.
- [4] Chen, H., Zhang, Y., Zhang, Q., et al. Maybe only 0.5% data is needed: A preliminary exploration of low training data instruction tuning. ArXiv, 2023.
- [5] Lu, K., Yuan, H., Yuan, Z., et al. #instag: Instruction tagging for analyzing supervised fine-tuning of large language models. ICLR, 2024.
- [6] Wu, S., Lu, K., Xu, B., et al. Self-evolved diverse data sampling for efficient instruction tuning. ArXiv, 2023.
- [7] Maharana, A., Yadav, P., Bansal, M. D² pruning: Message passing for balancing diversity & difficulty in data pruning. ICLR, 2024.
- [8] Zhou, D., Wang, K., Gu, J., et al. Dataset quantization. ICCV, 2023.
- [9] Köpf, A., Kilcher, Y., von Werra, L., et al. OpenAssistant Conversations–Democratizing Large Language Model Alignment. ArXiv, 2023.
- [10] Xia, M., Malladi, S., Gururangan, S., et al. Less: Selecting influential data for targeted instruction tuning. ArXiv, 2024.
- [11] Luo, J., Luo, X., Ding, K., et al. RobustFT: Robust Supervised Fine-tuning for Large Language Models under Noisy Response. ArXiv, 2024.
- [12] Zhang, D., Dai, Q., Peng, H. The Best Instruction-Tuning Data are Those That Fit. ArXiv, 2025.
- [13] Eldan, R., Mirhoseini, A., Norouzi, M. Who's Harry Potter? Approximate Unlearning in LLMs. ArXiv, 2023.

- [14] Jin, Z., Cao, P., Wang, C., et al. RWKU: Benchmarking real-world knowledge unlearning for large language models. ArXiv, 2024.
- [15] De Cao, N., Aziz, W., Titov, I. Editing Factual Knowledge in Language Models. EMNLP, 2021.
- [16] Jang, J., Yoon, D., Yang, S., et al. Knowledge unlearning for mitigating privacy risks in language models. ArXiv, 2022.
- [17] Maini, P., Feng, Z., Schwarzschild, A., et al. Tofu: A task of fictitious unlearning for llms. ArXiv, 2024.
- [18] Yao, Y., Xu, X., Liu, Y. Large language model unlearning. NeurIPS, 2024.
- [19] Tian, B., Liang, X., Cheng, S., et al. To forget or not? towards practical knowledge unlearning for large language models. ArXiv, 2024.
- [20] Cha, S., Cho, S., Hwang, D., Lee, M. Towards robust and cost-efficient knowledge unlearning for large language models. ArXiv, 2024.
- [21] Chen, L., Wang, Z., Zhang, H., et al. Unlearning Bias in Language Models by Partitioning Gradients. ArXiv, 2023.
- [22] Wang, Q., Han, B., Yang, P., et al. Towards Effective Evaluations and Comparisons for LLM Unlearning Methods. ICLR, 2024.
- [23] Wang, Q., Han, B., Yang, P., et al. Unlearning with control: Assessing real-world utility for large language model unlearning. ArXiv, 2024.
- [24] Lizzo, T., Heck, L. Unlearn: Efficient removal of knowledge in large language models. ArXiv, 2024.
- [25] Zhang, R., Lin, L., Bai, Y., Mei, S. Negative Preference Optimization: From Catastrophic Collapse to Effective Unlearning. ArXiv, 2024.
- [26] Yao, J., Chien, E., Du, M., et al. Machine unlearning of pre-trained large language models. ACL, 2024.
- [27] Wang, Q., Zhou, J.P., Zhou, Z., et al. Rethinking llm unlearning objectives: A gradient perspective and go beyond. ArXiv, 2025.
- [28] Schulman, J., Wolski, F., Dhariwal, P., et al. Proximal Policy Optimization Algorithms. ArXiv, 2017.
- [29] Rafailov, R., Sharma, A., Mitchell, E., et al. Direct preference optimization: Your language model is secretly a reward model. NeurIPS, 2023.
- [30] Ethayarajh, K., Xu, W., Muennighoff, N., et al. KTO: Model alignment as prospect theoretic optimization. ICLR, 2024.

- [31] Azar, M.G., Rowland, M., Piot, B., et al. A general theoretical paradigm to understand learning from human preferences. ArXiv, 2023.
- [32] Xu, H., Sharaf, A., Chen, Y., et al. Contrastive preference optimization: Pushing the boundaries of LLM performance in machine translation. ACL, 2024.
- [33] Hong, J., Lee, N., Thorne, J. ORPO: Monolithic preference optimization without reference model. ArXiv, 2024.
- [34] Meng, Y., Xia, M., Chen, D. SimPO: Simple preference optimization with a reference-free reward. ArXiv, 2024.
- [35] Zeng, Y., Liu, G., Ma, W., et al. Token-level direct preference optimization. ArXiv, 2024.
- [36] Zhang, R., Lin, L., Bai, Y., Mei, S. Negative Preference Optimization: From Catastrophic Collapse to Effective Unlearning. ArXiv, 2024.
- [37] Taheri Ghahrizjani, A., Taban, A., Ye, S., Mirzaei, A., Liu, T., Han, B. Forgetting: A New Mechanism Towards Better Large Language Model Fine-tuning. ArXiv, 2025.