

گیلان

- مفهوم SDN و کاربردهای آن
- کاربرد شبکه‌های عصبی در پردازش زبان طبیعی (NLP)
- مدل‌های زبانی بزرگ (LLMs)
- و ...

گیلانہ

بدون رشد و پیشرفت مستمر، کلماتی مانند پیشرفت، دستاورد و موفقیت معنایی ندارند.

بنجامین فرانکلین

ماہنامہ علمی دانشجویی - سال سوم - شماره چہل و دوم - ۲۹ بہمن ۱۴۰۳

صاحب امتیاز: انجمن علمی مہندسی کامپیوتر دانشگاه آزاد اسلامی واحد لاهیجان

مدیر مسئول: سورنا کریمی سلیمی

سردبیر: ریحانہ محمدپور

ویراستار: فرہاد فخری، سعید نیکفر

طراحی جلد و صفحہ آرایہ: معصومہ چنانہن

ہیأت تحریرہ:

حجت آزادروش، ملیکا آذین مہر، پارسا اپک چیان عسکری، امیرحسین شبرنگ، امیررضا دانشور، سعید رفعتی، بہنام قہرمان خانی، سروش قلیزادہ، امیررضا پورکریم، محمود علیمزادی، علی اشرف، مجتبی پورطہماسبی، دانیال غفاری، سینا خیراندیش، کمیل آقابابایی، محمد طالبی، محمد علیزادہ، سروش سہرابی، امیر اسدنیا، محمدصادق پاکزاد، مہتا مہرگانی، آیرین پورعباس، پویان قنبری، آیدا میرزازادہ، حوری طالش امیری، زہرا قنبریور، فاطمہ گل محمدی، محمد امید پور

فهرس

۴

خرده‌نان دیجیتال، راهنمای کاربران گم‌شده



۱۰

پیش‌گیری از خطرات دیجیتالی شدن؛ رویکردی روان‌شناختی برای دنیای دیجیتال امن‌تر



۱۳

DeepSeek چیست و چه کارایی‌هایی دارد؟



۱۵

مفهوم SDN و کاربردهای آن



۱۸

چگونگی بهره‌برداری از مدل‌های زبانی بزرگ، از دستوردهی تا یادگیری در متن



۲۳

کاربرد شبکه‌های عصبی در پردازش زبان طبیعی (NLP)



۲۷

مدل‌های زبانی بزرگ (LLMs)



۳۱

مدل‌های زبانی نقاب‌دار (Masked Language Models)



۳۴

شبکه‌های عصبی بازگشتی (RNNs) و حافظه طولانی کوتاه‌مدت (LSTMs)





قسمت بالایی صفحه به‌صورت افقی و با استفاده از جداکننده‌هایی مانند / یا > نمایش داده می‌شود.



تأثیر بردکرامب بر سئوی سایت

با استفاده از بردکرامب و انجام بهینه‌سازی‌های لازم برای سئو، می‌توانید تجربه‌ی کاربری و عملکرد سایتتان در نتایج جستجوی گوگل را بهبود دهید. هنگامی که بردکرامب‌ها به‌درستی پیاده‌سازی شوند، گوگل آن‌ها را به‌صورت یک نمای خلاصه و منظم معروف به ریح‌اسنیپت (Rich Snippet) در نتایج جست‌وجو نمایش می‌دهد. این نمایش غالباً شامل اطلاعاتی مانند امتیازات کاربران، تعداد نظرات و دیگر داده‌های مفید است که به کاربران کمک می‌کند تا راحت‌تر تصمیم به کلیک بگیرند و این امر نرخ کلیک (CTR) را افزایش می‌دهد.

اما مزایای بردکرامب فقط به ریح‌اسنیپت محدود نمی‌شود. این ابزار باعث بهبود ساختار سایت برای موتورهای جست‌وجو می‌شود و با ارائه‌ی یک مسیر واضح، تجربه‌ی کاربری بهتری ایجاد می‌کند. همچنین، کاربران می‌توانند با سرعت بیشتری به بخش‌های مختلف سایت دسترسی داشته باشند، که این موضوع نه تنها تعامل آن‌ها را افزایش می‌دهد، بلکه به کاهش نرخ پرش (Bounce Rate) نیز کمک می‌کند. در نتیجه، بردکرامب‌ها یکی از عناصر کلیدی در بهینه‌سازی سایت برای موتورهای جست‌وجو و افزایش ترافیک هدفمند به شمار می‌روند.

در دنیای وب، تجربه‌ی کاربری (UX) و دسترسی آسان به محتوای سایت از جمله عوامل کلیدی موفقیت هر وب‌سایتی محسوب می‌شوند. کاربران اغلب به دنبال روشی ساده و کارآمد برای حرکت در میان صفحات یک سایت هستند تا سریع‌تر به اطلاعات مورد نیازشان دست یابند. اینجاست که ابزارهای مسیریابی مانند بردکرامب (Breadcrumb) که به آن خرده‌نان نیز گفته می‌شود، نقش بسیار مهمی ایفا می‌کند. بردکرامب نه تنها به کاربران کمک می‌کند تا موقعیت خود را در ساختار سایت پیدا کنند، بلکه به موتورهای جست‌وجو نیز در درک بهتر ساختار سایت کمک می‌کند.

بردکرامب به‌عنوان یکی از ابزارهای پرکاربرد در طراحی وب‌سایت، می‌تواند تجربه‌ی کاربری را بهبود بخشد و سئوی سایت را تقویت کند. اما این ابزار چگونه عمل می‌کند؟ چرا برای سایت شما ضروری است؟ و چطور می‌توان آن را به‌طور مؤثر طراحی و پیاده‌سازی کرد؟ در این مقاله به بررسی مفاهیمی از قبیل تعریف بردکرامب، تأثیر آن بر سئوی سایت، مزایای استفاده از این ابزار، انواع بردکرامب، روش‌های اضافه کردن آن به وب‌سایت و در نهایت ۱۰ نکته‌ی کلیدی برای طراحی بردکرامب پرداخته شده‌است. با مطالعه‌ی این مقاله، با تمام جنبه‌های بردکرامب و نحوه‌ی استفاده مؤثر از آن در وب‌سایت خود آشنا خواهید شد.

بردکرامب چیست؟

بردکرامب (Breadcrumb) به معنی «خرده‌نان» است و به مجموعه‌ای از لینک‌ها اطلاق می‌شود که به کاربر مسیر حرکتش در یک وب‌سایت را نشان می‌دهند. این لینک‌ها به کاربران کمک می‌کنند تا به‌طور دقیق بدانند که در کدام بخش از سایت قرار دارند و چگونه به آنجا رسیدند. درحقیقت، بردکرامب نوعی نقشه‌راه است که باعث سهولت ناوبری می‌شود. این سیستم معمولاً در

مزایای استفاده از بردکرامب

حال که با مفهوم بردکرامپ آشنا شدیم در ادامه به برخی از مهم‌ترین مزایای استفاده از این قابلیت اشاره شده‌است:

۱- افزایش تمایل کاربران به گشت‌وگذار و بررسی در سایت

بردکرامب‌ها کاربران را تشویق می‌کنند تا از صفحات بیشتری در سایت بازدید کنند. این موضوع می‌تواند به افزایش نرخ تبدیل و حتی فروش منجر شود. برای مثال، فرض کنید کاربری به صفحه‌ی محصولات گوشی سامسونگ وارد می‌شود اما محصول موردنظرش را پیدا نمی‌کند. به‌جای ترک سایت، با دیدن مسیرهای لینک‌شده در بردکرامب، تصمیم می‌گیرد به بخش دسته‌بندی‌ها بازگردد و دیگر محصولات گوشی سامسونگ را بررسی کند.

۲- کمک به ربات‌های گوگل برای درک بهتر سایت

بردکرامب تنها برای کاربران مفید نیست، بلکه نقش مهمی در سئو دارد. این ابزار به موتورهای جست‌وجو کمک می‌کند تا ساختار سایت را بهتر و سریع‌تر درک کنند. بردکرامب مانند نقشه‌ای شفاف، ارتباط بین صفحات مختلف سایت را به گوگل نشان می‌دهد. این کار باعث می‌شود گوگل صفحات شما را راحت‌تر پیدا کند و احتمال نمایش آن‌ها در نتایج جست‌وجو افزایش یابد. درواقع، بردکرامب یک ابزار قدرتمند برای بهبود رتبه‌ی سایت در کلمات کلیدی محسوب می‌شود.

۳- بهبود تجربه‌ی کاربری و جلوگیری از سردرگمی

تا حالا شده در یک سایت، از صفحه‌ای به صفحه‌ی دیگر منتقل شوید و دیگر ندانید چگونه به آنجا رسیده‌اید؟ در این مواقع، معمولاً کاربران مجبور می‌شوند چندین بار دکمه برگشت مرورگر را فشار دهند تا به صفحه اول بازگردند. بردکرامب این مشکل را حل می‌کند. با این ابزار، کاربران دقیقاً می‌دانند کجای سایت قرار دارند و تنها با یک کلیک می‌توانند به بخش‌های قبلی یا اصلی سایت برگردند.

۴- کاهش نرخ پرش (Bounce Rate)

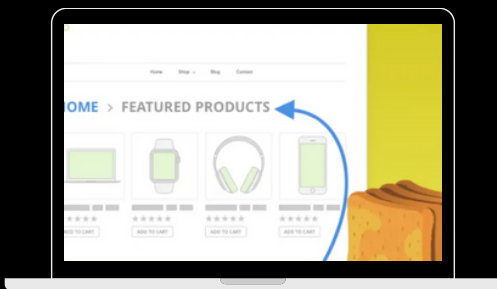
یکی از دلایل اصلی خروج کاربران از سایت‌ها، سردرگمی و ناتوانی در یافتن اطلاعات موردنظرشان است. بردکرامب با ارائه‌ی یک مسیر مشخص و آسان، به کاربران کمک می‌کند تا به‌جای ترک سایت، مسیرهای دیگر را برای یافتن اطلاعات یا محصولات بررسی کنند. این موضوع مستقیماً به کاهش نرخ پرش و افزایش تعامل کاربران با سایت منجر می‌شود.

۵- بهبود دسترسی‌پذیری (Accessibility)

بردکرامب یک ابزار کلیدی برای بهبود دسترسی‌پذیری سایت محسوب می‌شود. این قابلیت به‌ویژه برای کاربران با محدودیت‌های حرکتی یا افرادی که از ابزارهای کمکی استفاده می‌کنند، مفید است. بردکرامب با ساده کردن مسیرها و نمایش دقیق مکان کاربر در سایت، تجربه‌ی بهتری برای این گروه از کاربران فراهم می‌کند.

۶- تسهیل مدیریت و ویرایش سایت

بردکرامب‌ها نه‌تنها برای کاربران و موتورهای جست‌وجو مفید هستند، بلکه مدیریت سایت را برای صاحبان آن ساده‌تر می‌کنند. این ابزار به مدیران سایت کمک می‌کند تا ساختار و ارتباط بین صفحات مختلف را بهتر درک کنند و در صورت نیاز، تغییرات لازم را به‌سادگی اعمال کنند. همچنین، بردکرامب‌ها ابزاری مفید برای تحلیل رفتار کاربران در سایت هستند و می‌توانند به بهبود استراتژی‌های محتوا و طراحی کمک کنند.

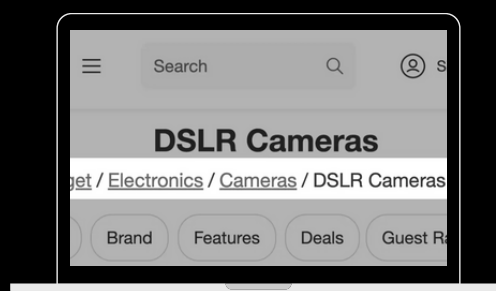


انواع بردکرامب

۱- بردکرامب مبتنی بر سلسله‌مراتب (Hierarchy-Based Breadcrumbs)

این نوع بردکرامب متداول‌ترین مدل است و بر اساس ساختار سلسله‌مراتبی وب‌سایت عمل می‌کند. برای مثال، اگر کاربری وارد دسته‌بندی «عکاسی» شود، بردکرامب مبتنی بر سلسله‌مراتب مسیر را به صورت زیر نمایش می‌دهد:

صفحه اصلی < عکاسی < دوربین‌های DSLR
این نوع بردکرامب به‌ویژه برای کاربرانی که از طریق موتورهای جستجو به یک صفحه‌ی خاص وارد می‌شوند بسیار کاربردی است. چرا که به آن‌ها امکان می‌دهد به آسانی به دسته‌بندی‌های بالاتر برگردند و سایر بخش‌های وب‌سایت را نیز بررسی کنند. این کار نه تنها تجربه‌ی کاربری را بهبود می‌بخشد، بلکه می‌تواند باعث افزایش تعامل کاربران و کاهش نرخ پرش شود.



۲- بردکرامب مبتنی بر ویژگی (Attribute-Based Breadcrumbs)

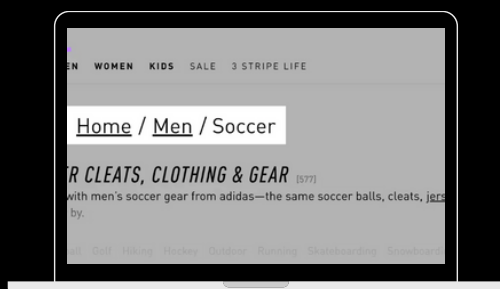
این نوع بردکرامب بر اساس فیلترهایی که کاربران در سایت اعمال می‌کنند، مسیر را به صورت پویا تغییر می‌دهد. بردکرامب مبتنی بر ویژگی معمولاً در فروشگاه‌های اینترنتی استفاده می‌شود و اغلب در کنار بردکرامب مبتنی بر سلسله‌مراتب به کار می‌رود.

برای مثال، فرض کنید وارد یک سایت فروش لوازم ورزشی شده‌اید و فیلتر «برند» را روی «آدیداس» اعمال کرده‌اید. سپس وارد بخش «کفش‌های دویدن» می‌شوید. حالا بردکرامب مبتنی بر ویژگی این مسیر را به شما نشان می‌دهد و این امکان را فراهم می‌کند که

به راحتی سایر کفش‌های دویدن از برند آدیداس را مشاهده کنید.

۳- بردکرامب مبتنی بر تاریخچه (History-Based Breadcrumbs)

این نوع بردکرامب، مسیر حرکت کاربر را تا رسیدن به صفحه‌ی فعلی نمایش می‌دهد. برخلاف بردکرامب‌های سلسله‌مراتبی یا مبتنی بر ویژگی، این مدل به جای نمایش ساختار سایت، بر تاریخچه‌ی حرکت کاربر تمرکز دارد. معمولاً به جای نمایش جزئیات کامل مسیر، تنها یک گزینه برای بازگشت به صفحه‌ی قبلی ارائه می‌دهد. با این حال، به دلیل وجود دکمه‌ی «بازگشت» در مرورگرها، استفاده از بردکرامب مبتنی بر تاریخچه چندان رایج نیست. طراحان اغلب ترجیح می‌دهند از مدل‌هایی استفاده کنند که علاوه بر سادگی، قابلیت‌های بیشتری مانند ناوبری میان بخش‌های مختلف سایت را نیز فراهم کنند. با این حال، این نوع بردکرامب ممکن است در برخی سناریوهای خاص، مانند برنامه‌های تحت وب یا اپلیکیشن‌های پیچیده، که مسیر حرکت کاربر اهمیت ویژه‌ای دارد، مورد استفاده قرار گیرد.



چگونه بردکرامب را به سایت اضافه کنیم؟

برای افزودن بردکرامب به وب‌سایت، دو روش کلی پیش روی شماست:

۱- استفاده از کد HTML

۲- استفاده از افزونه‌ها

اضافه کردن بردکرامب به وردپرس با کد HTML

ساده‌ترین روش برای پیاده‌سازی بردکرامب در سایت، استفاده از کد HTML است. کافی است لینک‌های مسیر را در قالب یک لیست بدون ترتیب () سازمان‌دهی کنید. هر آیتم لیست () باید شامل یک لینک باشد که نمایانگر هر مرحله از مسیر بردکرامب تا رسیدن به صفحه‌ی فعلی است. آخرین آیتم، صفحه‌ای را نشان می‌دهد که کاربر در آن قرار دارد و معمولاً به‌صورت غیرلینک نمایش داده می‌شود.

در زیر، یک نمونه الگوی HTML ارائه شده که می‌توانید از آن به‌عنوان نقطه‌ی شروع استفاده کنید:

```
<nav aria-label="breadcrumb">
  <ul class="breadcrumb">
    <li><a href="/">صفحه اصلی</a></li>
    <li><a href="/category">دسته‌بندی</a></li>
    <li class="active">صفحه فعلی</li>
  </ul>
</nav>
```

اضافه کردن بردکرامب به وردپرس با افزونه

وردپرس، به‌عنوان محبوب‌ترین سیستم مدیریت محتوا (CMS)، امکانات گسترده‌ای برای بهبود تجربه‌ی کاربری و افزایش رتبه‌ی سئو فراهم می‌کند. یکی از این ابزارها، مسیر راهنما یا همان بردکرامب است که کمک شایانی به مسیریابی کاربران می‌کند. بسیاری از قالب‌های وردپرس به‌صورت پیش‌فرض از بردکرامب پشتیبانی می‌کنند اما اگر قالب شما این ویژگی را ندارد، نگران نباشید! با استفاده از افزونه‌های بردکرامب، می‌توانید به‌راحتی این قابلیت را به وبسایت خود اضافه کنید.

یکی از بهترین افزونه‌ها برای افزودن بردکرامب، Yoast SEO است. این افزونه نه‌تنها یک ابزار حرفه‌ای برای بهینه‌سازی سایت شما در موتورهای جست‌وجو است، بلکه قابلیت اضافه کردن بردکرامب را نیز فراهم می‌کند.

از دیگر افزونه‌های محبوب بردکرامب در وردپرس می‌توان به افزونه‌های RankMath، BreadCrumb، و All In One SEO اشاره کرد.

۱۰ نکته‌ی کلیدی برای طراحی بردکرامب در سایت

بردکرامب ابزار مفیدی است که به کاربران کمک می‌کند موقعیت خود را در ساختار سایت تشخیص دهند و به‌راحتی بین صفحات مختلف حرکت کنند. در ادامه، به نکات مهمی اشاره می‌کنیم که هنگام طراحی و استفاده از بردکرامب باید در نظر داشته باشید:

۱- فقط زمانی از بردکرامب استفاده کنید که ساختار سایت پیچیده باشد.

اگر سایت شما تنها شامل چند صفحه‌ی ساده است و کاربران می‌توانند از طریق منوی اصلی به تمام بخش‌ها دسترسی پیدا کنند، نیازی به استفاده از بردکرامب نیست. اما اگر وبسایت شما صفحات متعددی دارد و کاربران باید میان دسته‌بندی‌ها یا مسیرهای مختلف جابه‌جا شوند، استفاده از بردکرامب می‌تواند تجربه‌ی کاربری را بهبود بخشد.

۲- بردکرامب را سبک و کم‌حجم طراحی کنید.

بردکرامب نقش ثانویه‌ای نسبت به منوی اصلی سایت دارد. بنابراین بهتر است آن را در اندازه‌ای کوچک‌تر و با استایلی ساده‌تر طراحی کنید تا ظاهر کلی سایت شلوغ و گیج‌کننده نشود. قرار دادن بردکرامب در محلی پایین‌تر از منوی اصلی نیز به حفظ تعادل بصری سایت کمک می‌کند.

۳- مسیر کامل را در بردکرامب نشان دهید.

بسیاری از کاربران از طریق نتایج جست‌وجو مستقیم وارد صفحات داخلی سایت می‌شوند، نه صفحه‌ی اصلی. به همین دلیل، ضروری است مسیر کامل از صفحه‌ی اصلی تا صفحه‌ی فعلی در بردکرامب نمایش داده شود. این کار به کاربران کمک می‌کند تا حتی اگر از مسیر اصلی وارد سایت نشده باشند، به‌آسانی موقعیت خود را در ساختار سایت پیدا کرده و در صورت نیاز، به دسته‌بندی‌ها یا صفحات قبلی بازگردند.

۴- مسیر حرکت را از بالا به پایین و منطقی نمایش دهید.

در طراحی بردکرامب، ترتیب نمایش لینک‌ها اهمیت بسیاری دارد. ساختار باید به گونه‌ای باشد که اولین لینک، صفحه‌ی اصلی و آخرین لینک، صفحه‌ی فعلی باشد. تحقیقات نشان داده‌اند کاربران برای زبان‌های راست‌چین مانند فارسی، بیشتر به سمت راست صفحه توجه می‌کنند. بنابراین در سایت‌های فارسی، بهتر است بردکرامب در سمت راست و در سایت‌های چپ‌چین مانند انگلیسی، در سمت چپ نمایش داده شود. این رویکرد با عادت‌های کاربران تطابق دارد و تجربه‌ی کاربری را بهبود می‌بخشد.

۵- هماهنگی بین عنوان‌های بردکرامب و صفحات

برای جلوگیری از سردرگمی کاربران، حتماً عناوین بردکرامب را با عنوان‌های صفحات تطابق دهید. این هماهنگی نه تنها مسیر حرکت را شفاف‌تر می‌کند، بلکه به شما فرصت می‌دهد کلمات کلیدی مرتبط را دوباره به کاربران و موتورهای جستجو یادآوری کنید. همچنین، حتماً مشخص کنید کدام بخش از بردکرامب قابل کلیک است و کدام یک فقط نمایش مسیر را بر عهده دارد تا کاربر به درستی هدایت شود.

۶- نوآوری در طراحی بردکرامب

بردکرامب اغلب به صورت لیستی افقی با جداکننده‌هایی مانند «>» طراحی می‌شود، اما شما محدود به این الگو نیستید. می‌توانید از علامت‌های دیگر مانند «/» یا حتی طراحی‌های سفارشی و گرافیکی استفاده کنید. البته این خلاقیت باید با طراحی کلی سایت هماهنگ باشد تا تجربه‌ی کاربری تحت تأثیر قرار نگیرد.

۷- تمیزی و سادگی، اصل اساسی

بردکرامب باید به کاربران کمک کند، نه اینکه آن‌ها را گیج کند یا حواسشان را پرت کند. بنابراین طراحی بردکرامب باید ساده، تمیز و با متن‌های مختصر باشد. از شلوغ‌کاری در طراحی یا اضافه کردن اطلاعات غیرضروری پرهیز کنید تا کاربران بتوانند به راحتی مسیر خود را پیدا کنند.

۸- انتخاب نوع بردکرامب متناسب با نیاز سایت

بردکرامب‌ها در سه نوع مکان‌محور، ویژگی‌محور و تاریخچه‌محور ارائه می‌شوند.

الف) مکان‌محور: موقعیت صفحه‌ی فعلی در ساختار کلی سایت را نشان می‌دهد.

ب) ویژگی‌محور: بر اساس فیلترها یا ویژگی‌هایی که کاربر انتخاب کرده، مسیر را نمایش می‌دهد.

ج) تاریخچه‌محور: مسیر واقعی طی‌شده توسط کاربر را نشان می‌دهد.

برای انتخاب نوع مناسب، ابتدا ساختار سایت خود و نیازهای کاربران را بررسی کنید.

۹- شناخت مخاطب و تنظیم موقعیت بردکرامب

محل قرارگیری بردکرامب باید با عادت‌های مخاطبان شما هماهنگ باشد. معمولاً بردکرامب بالای صفحه و پایین‌تر از منوی اصلی قرار می‌گیرد. با این حال، برخی سایت‌ها مانند اپل بردکرامب را در پایین صفحه و داخل فوتر قرار می‌دهند؛ زیرا مخاطبان آن‌ها اغلب کاربران حرفه‌ای هستند که به راحتی مسیر را پیدا می‌کنند. از ابزارهایی مانند (A/B Test) استفاده کنید تا بهترین محل قرارگیری بردکرامب را بر اساس رفتار کاربران تعیین کنید.

۱۰- از میکرودیتا (Schema Markup) برای بهبود سئو استفاده کنید

یکی از نکات پیشرفته برای بردکرامب، استفاده از میکرودیتا یا اسکیمای مارکاپ است. این تکنیک به موتورهای جستجو کمک می‌کند ساختار بردکرامب سایت شما را بهتر درک کنند و آن را به صورت جذاب‌تری در نتایج جستجو نمایش دهند. برای این کار می‌توانید از استانداردهای اسکیمای مثل BreadcrumbList استفاده کنید. این کار نه تنها تجربه‌ی کاربری را بهبود می‌بخشد، بلکه احتمال کلیک کاربران بر لینک سایت شما در نتایج جستجو را افزایش می‌دهد.

سخن آخر

بردگرماب یکی از ابزارهای ساده اما بسیار مؤثر در بهبود تجربه‌ی کاربری و ارتقای سئوی سایت است. این ابزار مسیریابی با ایجاد راهنمایی شفاف برای کاربران، آن‌ها را قادر می‌سازد به راحتی در میان صفحات مختلف سایت حرکت کنند. از طرف دیگر، بردگرماب به موتورهای جستجو کمک می‌کند تا ساختار سایت را بهتر درک کنند و در نتیجه، شانس شما برای دیده شدن در نتایج جستجو افزایش پیدا می‌کند.

برای استفاده‌ی بهینه از بردگرماب، باید نکاتی مانند حفظ سادگی، هماهنگی با طراحی سایت، نمایش مسیر از بالا به پایین، و تطبیق عنوان‌ها با محتوای صفحات را در نظر بگیرید. علاوه بر این، شناخت مخاطبان و تست روش‌های مختلف می‌تواند به شما کمک کند که بهترین مکان و ساختار بردگرماب را انتخاب کنید.

در نهایت، بردگرماب نه تنها یک ابزار کاربردی برای کاربران و موتورهای جستجو است، بلکه فرصتی برای تقویت برند، افزایش نرخ تبدیل، و بهبود تجربه‌ی کلی کاربری به شمار می‌رود. با پیاده‌سازی اصولی این ابزار، می‌توانید مسیری شفاف و کارآمد برای کاربران خود ایجاد کرده و ارزش سایت خود را چندین برابر کنید.



اصطلاح اضطراب فومو (FOMO) یا (Fear of Missing) شناخته می‌شود که به ترس از دست دادن اطلاعات، ارتباطات یا فرصت‌ها اشاره دارد. این پدیده می‌تواند فرد را از زندگی واقعی خود دور کرده و به آن آسیب‌های روانی وارد کند.

۲- اعتیاد به فناوری و رسانه‌های اجتماعی

اعتیاد به فناوری و شبکه‌های اجتماعی یکی از بزرگ‌ترین چالش‌های روان‌شناختی ناشی از دیجیتالی شدن است. تحقیقات نشان می‌دهند که استفاده‌ی بیش‌ازحد از شبکه‌های اجتماعی می‌تواند به اختلالات روانی مانند افسردگی، اضطراب اجتماعی، و احساس تنهایی منجر شود. فضای دیجیتال به افراد این امکان را می‌دهد که به راحتی با دیگران ارتباط برقرار کنند. این وابستگی به فناوری می‌تواند موجب کاهش تمرکز، کاهش بهره‌وری در زندگی روزمره و حتی اختلالات خواب شود.

۳- اختلال در روابط اجتماعی و خانوادگی

در دنیای دیجیتال، به‌طور فزاینده‌ای شاهد افت روابط اجتماعی واقعی و ناتوانی افراد در برقراری ارتباطات عمیق و واقعی هستیم. این امر به‌ویژه در میان جوانان و نوجوانان قابل توجه است. نوجوانانی که بیشتر وقت خود را در فضای دیجیتال می‌گذرانند، ممکن است دچار مشکلاتی در مهارت‌های ارتباطی خود شوند و در برقراری ارتباط چهره‌به‌چهره و ایجاد روابط واقعی دچار مشکل شوند. علاوه بر این، استفاده‌ی مفرط از رسانه‌های اجتماعی و پیام‌رسان‌ها می‌تواند موجب بی‌توجهی به نیازهای عاطفی و روانی اعضای خانواده و دوستان شود. افراد ممکن است توجه خود را به دنیای آنلاین معطوف کنند و روابط خانوادگی و دوستانه‌ی خود را نادیده بگیرند که این امر می‌تواند به انزوای اجتماعی و احساس بی‌پناهی منجر شود.

در عصر حاضر، دنیای دیجیتال به یکی از اجزای جدایی‌ناپذیر زندگی روزمره تبدیل شده است. از ارتباطات روزمره و خریدهای آنلاین گرفته تا آموزش‌های مجازی و کارهای حرفه‌ای، همه چیز تحت تأثیر فناوری‌های دیجیتال قرار دارد. این روند دیجیتالی شدن هر روز بیشتر گسترش می‌یابد و به شکل عمیقی در بافت اجتماعی و فردی ما نفوذ می‌کند. با این حال، استفاده‌ی مفرط از دنیای دیجیتال می‌تواند با خود خطراتی به‌همراه داشته باشد که اغلب به جنبه‌های روان‌شناختی ما آسیب می‌زند. در این مطلب، به بررسی تهدیدات روان‌شناختی ناشی از دیجیتالی شدن پرداخته و راهکارهایی برای پیش‌گیری از این خطرات ارائه خواهیم داد.



تهدیدات روان‌شناختی دیجیتالی شدن

۱- افزایش اضطراب و استرس دیجیتال

استفاده‌ی زیاد از دستگاه‌های دیجیتال و پلتفرم‌های آنلاین می‌تواند موجب بروز اضطراب و استرس شود. این اضطراب معمولاً به دلیل اطلاعات و اخبار بی‌پایانی است که از طریق رسانه‌های اجتماعی و سایت‌ها به فرد منتقل می‌شود. پویایی‌های پیچیده‌ای که در این فضاها وجود دارند، می‌توانند به فرد احساس ناتوانی و غرق شدن در حجم اطلاعات و رویدادهای مختلف دهند. بسیاری از افراد تحت تأثیر این اضطراب‌ها احساس فشار می‌کنند تا همواره آنلاین باشند، اخبار را پیگیری کنند و در فضای مجازی حاضر و فعال باشند. این نوع اضطراب معمولاً

۴- اختلالات هویتی و فشارهای اجتماعی دیجیتال

یکی دیگر از تهدیدات روان‌شناختی دیجیتالی شدن، ایجاد اختلالات هویتی و فشارهای اجتماعی است که به‌طور خاص در نوجوانان و جوانان مشاهده می‌شود. فضای دیجیتال با ترویج تصاویر ایده‌آل و استانداردهای زیبایی و زندگی، موجب فشار روانی زیادی به افراد می‌شود. این فشار می‌تواند به خلق‌و‌خو، اعتماد به نفس و احساسات درونی فرد آسیب بزند. این پدیده را می‌توان از دیدگاه روان‌شناسی اجتماعی به‌عنوان یک مشکل هویت اجتماعی توضیح داد. فضای مجازی، به‌ویژه شبکه‌های اجتماعی، محیطی را فراهم می‌آورد که افراد به‌طور مداوم خود را با دیگران مقایسه می‌کنند و تلاش می‌کنند تا تصویری از خود ارائه دهند که مورد تأیید جمع قرار گیرد. این روند می‌تواند به کاهش عزت نفس، اضطراب اجتماعی و حتی اختلالات افسردگی منجر شود.



راهکارهای پیش‌گیری از خطرات روان‌شناختی دیجیتالی شدن

۱- آموزش خودآگاهی و مدیریت استرس دیجیتال

یکی از راه‌های پیش‌گیری از اضطراب و استرس دیجیتال، یادگیری خودآگاهی و مدیریت استرس است. برای مثال، تکنیک‌های مدیتیشن، تنفس عمیق، و تمرین‌های ذهن‌آگاهی می‌توانند به کاهش اضطراب ناشی از استفاده مفرط از فناوری کمک کنند. این تکنیک‌ها به افراد کمک می‌کنند تا ذهن خود را از اطلاعات بیش از حد و فشار روانی ناشی از فضای

دیجیتال پاک کنند و به آرامش دست یابند. مدیریت زمان در فضای دیجیتال نیز از اهمیت ویژه‌ای برخوردار است. افراد باید خود را متعهد کنند که زمان‌های مشخصی را به استفاده از اینترنت و شبکه‌های اجتماعی اختصاص دهند و از مرور بی‌وقفه‌ی محتواهای آنلاین جلوگیری کنند. برنامه‌ریزی دقیق و محدود کردن زمان استفاده از گوشی‌های هوشمند و رایانه‌ها می‌تواند به فرد کمک کند تا احساس بهتری در مورد حضور آنلاین خود داشته باشد.

۲- ایجاد محدودیت‌های استفاده از فناوری برای پیش‌گیری از اعتیاد

برای مقابله با اعتیاد به رسانه‌های اجتماعی، یکی از بهترین روش‌ها، تعیین محدودیت‌های زمانی و برنامه‌ریزی دقیق برای استفاده از این پلتفرم‌ها است. تحقیقات نشان می‌دهند که افرادی که از ابزارهای محدودکننده‌ی زمان استفاده می‌کنند، کمتر دچار آسیب‌های روانی ناشی از اعتیاد به فناوری می‌شوند. برای مثال، تنظیم محدودیت‌هایی برای زمان استفاده از گوشی‌های هوشمند، مشخص کردن زمان‌هایی برای حضور در شبکه‌های اجتماعی، و ایجاد زمان‌های بدون فناوری می‌تواند به فرد کمک کند تا تعادل بهتری در زندگی خود برقرار کند. توصیه‌ی دیگر برای کاهش آسیب‌های اعتیاد به رسانه‌های اجتماعی، پرهیز از چک کردن گوشی‌های هوشمند بلافاصله پس از بیدار شدن از خواب و پیش از خوابیدن است. این نوع رفتارهای خودآگاهانه به افراد کمک می‌کنند تا فضای دیجیتال را به‌عنوان ابزاری برای تسهیل زندگی خود استفاده کنند و نه به‌عنوان عاملی که بر سلامت روان آن‌ها تأثیر منفی می‌گذارد.

۳- تقویت مهارت‌های ارتباطی و اجتماعی در دنیای واقعی

برای مقابله با اثرات منفی استفاده مفرط از فضای دیجیتال بر روابط اجتماعی، تقویت مهارت‌های ارتباطی و

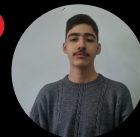
از بزرگ‌ترین چالش‌های روان‌شناختی عصر حاضر تبدیل شده‌اند. با این حال، با استفاده از رویکردهای پیش‌گیرانه مانند مدیریت استرس دیجیتال، محدودیت‌های زمانی در استفاده از فناوری، تقویت روابط اجتماعی واقعی و حفظ حریم خصوصی، می‌توان از خطرات روان‌شناختی دیجیتالی شدن جلوگیری کرد و به یک زندگی دیجیتال سالم و متعادل دست یافت.

اجتماعی در دنیای واقعی ضروری است. افراد باید از فضای دیجیتال فاصله بگیرند و زمان بیشتری را به ایجاد روابط نزدیک‌تر با خانواده و دوستان اختصاص دهند. یادگیری هنر گفت‌وگو و برقراری ارتباط چهره‌به‌چهره نیز به تقویت روابط اجتماعی کمک خواهد کرد. در این راستا، شرکت در فعالیت‌های اجتماعی، گروه‌های آموزشی، یا ورزش‌های گروهی می‌تواند فرصت‌های مناسبی برای ایجاد روابط اجتماعی واقعی و کاهش تأثیرات منفی دنیای دیجیتال بر فرد فراهم آورد. به‌طورخاص، ایجاد فضای حمایتی برای افراد آسیب‌دیده از اثرات روانی دیجیتالی شدن، نظیر جلسات مشاوره‌ی گروهی و روان‌درمانی آنلاین، می‌تواند به آن‌ها کمک کند تا از این چالش‌ها عبور کنند. همچنین، افراد باید یاد بگیرند که چگونه خود را از فشارهای اجتماعی دیجیتال رها کنند و به خود اجازه دهند که بدون نیاز به تأیید دیگران، هویت واقعی خود را پرورش دهند.

۴- حفظ حریم خصوصی و جلوگیری از آسیب‌های اجتماعی آنلاین

آگاهی از خطرات امنیتی و روان‌شناختی آنلاین و محافظت از حریم خصوصی شخصی می‌تواند به جلوگیری از مشکلات روان‌شناختی ناشی از نقض حریم خصوصی و سوءاستفاده‌های آنلاین کمک کند. این امر شامل استفاده از رمزهای عبور پیچیده، فعال‌سازی احراز هویت دو مرحله‌ای، و محدود کردن اشتراک‌گذاری اطلاعات شخصی در فضای آنلاین است. از منظر روان‌شناسی امنیت دیجیتال، این اقدامات می‌توانند احساس امنیت و کنترل بیشتری به افراد دهند و به جلوگیری از اضطراب و استرس ناشی از حملات سایبری یا سوءاستفاده از داده‌های شخصی کمک کنند.

دیجیتالی شدن در حالی که به زندگی انسان‌ها امکانات جدیدی می‌دهد، چالش‌ها و تهدیدات روان‌شناختی زیادی نیز به‌همراه دارد. از اضطراب دیجیتال و استرس اطلاعاتی گرفته تا اعتیاد به فناوری و اختلالات هویتی، مشکلات روانی ناشی از دیجیتالی شدن به‌تدریج به یکی



- (ب) مدل (R1):
- هدف اصلی: کاربردهای تخصصی و پیشرفته.
- پشتیبانی زبانی: بهینه‌سازی برای زبان‌های خاص.
- تحلیل داده‌ها: مناسب برای داده‌های پیچیده و تخصصی.
- سرعت و دقت: دقت بالاتر و دارای قابلیت تحلیلی پیشرفته.
- کاربردهای مالی: مناسب برای تحلیل‌های مالی پیشرفته و پیش‌بینی بازار.

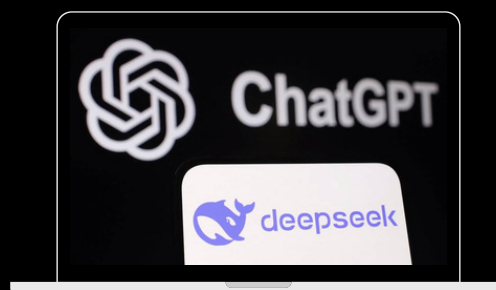
دیپ‌سیک (DeepSeek)، و تأثیر آن بر بازارهای کریپتو
 در ژانویه سال ۲۰۲۵ به دنبال افزایش توجه به موفقیت‌های دیپ‌سیک (DeepSeek) و رقابت آن با چت‌جی‌پی‌تی (ChatGPT) قیمت به‌طور قابل‌توجهی سقوط کرد. در کمتر از ۴۸ ساعت، قیمت بیت‌کوین از ۱۰۵۰۰۰ دلار به ۹۸۰۰۰ دلار کاهش یافت. تحلیلگران این تغییرات را ناشی از نگرانی‌ها نسبت به بی‌ثباتی سهام شرکت‌های هوش مصنوعی مانند NVIDIA و Meta می‌دانند که سرمایه‌گذاران را به خروج از بازارهای دیجیتال ترغیب نمودند.

کاربرد دیپ‌سیک (DeepSeek) در بازار ارز دیجیتال
 در اینجا به برخی از مهم‌ترین کاربردهای دیپ‌سیک در بازار ارز دیجیتال اشاره خواهیم کرد:
 ۱- تحلیل پیشرفتهی داده‌های بازار: دیپ‌سیک می‌تواند داده‌های بازار ارز دیجیتال را تحلیل کند. به‌عنوان مثال دیپ‌سیک توانست در سال ۲۰۲۴ همبستگی بین قیمت اتریوم و انتشار اخبار مربوط به مقررات اتحادیهی اروپا را شناسایی کند.

هوش مصنوعی (AI) که به یکی از مهم‌ترین ابزارهای زندگی تبدیل شده‌است، به‌عنوان یک مدل پیشرفتهی زبانی، جایگاه ویژه‌ای در دنیای تعاملات هوشمندانه دارد. در این مقاله سعی داریم توضیح (DeepSeek-V3) و کاربردهای آن و همچنین تفاوت‌های آن با هوش مصنوعی (ChatGPT) بپردازیم.

دیپ‌سیک (DeepSeek) چیست؟

دیپ‌سیک (DeepSeekAI) یک استارت‌آپ چینی در حال رشد می‌باشد که با هزینه‌ی کمتر از ۶ میلیون دلار توسعه یافته‌است که این رقم بسیار پایین‌تر از هزینه‌ی شرکت‌های آمریکایی است. برای مثال اوپن‌ای‌آی (OpenAI) برای آموزش مدل GPT-4 حدود ۱۰۰ میلیون دلار و حتی برای آموزش مدل‌های استدلال‌محور (o1) و (o3) رقم‌های بسیار بالاتر هزینه کرده‌است.



تفاوت‌های بین دو مدل (o1 و o3) و (R1):

- الف) مدل (o1 و o3):
- هدف اصلی: کاربردهای عمومی و گسترده
- پشتیبانی زبانی: چندزبانه بودن «عمومی»
- تحلیل داده: مناسب برای داده‌های عمومی
- سرعت و دقت: سریع و دقیق و برای کاربردهای عمومی
- کاربردهای مالی: محدود به تحلیل‌های عمومی

۴- کار با API: دیپسیک می‌تواند کدی بنویسد که داده‌ها را از یک API دریافت و پردازش کند و می‌تواند یک API ایجاد کند.

۵- بهینه‌سازی کد (Code Optimization): می‌تواند عملکرد کد را بهبود ببخشد یا آن را خواناتر جلوه دهد که شامل کاهش پیچیدگی کد و بهبود عملکرد می‌باشد.

به‌طورکلی هوش مصنوعی (DeepSeek) نسبت به هوش مصنوعی چت‌جی‌پی‌تی (ChatGPT) عملکرد بهتری در زمینه‌های چت‌بات (Chatbot)، بازار ارزهای دیجیتال، مدل‌های استدلالی مثل R1 و برنامه‌نویسی دارد و بر اساس نظرات عموم رابط کاربری ساده و بهتری نسبت به ChatGPT ارائه می‌دهد.

۲- بهینه‌سازی سبد سرمایه‌گذاری: این هوش مصنوعی با استفاده از تئوری مارکوویتز (Modern Portfolio Theory)، ترکیب بهینه‌ای از ارزها را پیشنهاد می‌دهد.

۳- پیش‌بینی ریسک‌های بزرگ: دیپسیک با تحلیل داده‌های اقتصادی-سیاسی می‌تواند بازار کریپتو را مدل‌سازی کند.



کاربرد دیپسیک (DeepSeek) در برنامه‌نویسی

دیپسیک به‌عنوان یک نوع مدل هوش مصنوعی می‌تواند ابزار مناسبی برای برنامه‌نویس‌ها در زمینه‌های کد زدن، دیباگ کردن، بهینه‌سازی و... که ما به ۵ مورد اینجا اشاره خواهیم کرد.

۱- دیباگ کردن: دیپسیک برای برطرف کردن باگ‌های یک کد می‌تواند مناسب باشد و خطاها را تشخیص داده و بسته به کد شما راه‌حل‌های متنوعی بدهد.

۲- تبدیل کد (Code Translation): در تبدیل کردن کد خود از یک زبان به زبان دیگر به‌عنوان مثال تبدیل کد پایتون به جاوا.

۳- کار با کتابخانه‌ها و فریم‌ورک‌ها: این هوش مصنوعی سهولت کار در استفاده از کتابخانه‌ها و فریم‌ورک‌ها را فراهم کرده که می‌تواند نمونه‌ی کدهایی برای استفاده از کتابخانه‌ها یا فریم‌ورک‌ها ارائه دهد.



۳- لایه‌ی برنامه (Application Layer): این لایه شامل برنامه‌ها و اپلیکیشن‌هایی است که از API های ارائه شده توسط کنترلر SDN برای مدیریت و پیکربندی شبکه استفاده می‌کنند. این برنامه‌ها می‌توانند شامل ابزارهای مدیریت شبکه، نظارت بر ترافیک، امنیت و تحلیل داده‌ها باشند.

ویژگی‌ها و مزایای SDN

۱- قابلیت برنامه‌ریزی و مدیریت مرکزی: مدیران شبکه می‌توانند تمام عملیات و پیکربندی‌ها را از یک مکان مرکزی انجام دهند و نیازی به دست‌کاری هر دستگاه شبکه به صورت مجزا نیست. تغییرات در پیکربندی شبکه سریع‌تر و راحت‌تر انجام می‌شود.

۲- انعطاف‌پذیری و مقیاس‌پذیری: به دلیل جدا بودن لایه‌های داده و کنترل، شبکه می‌تواند به سرعت و بدون نیاز به تغییرات فیزیکی در دستگاه‌ها یا سخت‌افزار، مقیاس‌پذیر شود. این انعطاف‌پذیری می‌تواند در مدیریت شبکه‌های بزرگ و پیچیده بسیار مفید باشد.

۳- بهینه‌سازی ترافیک و کارایی شبکه: به کمک کنترلرهای SDN، می‌توان ترافیک شبکه را به صورت دینامیک مدیریت کرده و از ازدحام در برخی بخش‌ها جلوگیری کرد. می‌توان از الگوریتم‌های هوشمند برای مسیریابی بهینه استفاده کرد.

۴- کاهش هزینه‌های سخت‌افزاری: به دلیل جدا بودن کنترل و مدیریت از سخت‌افزار، می‌توان از تجهیزات ساده‌تری استفاده کرد و هزینه‌های سخت‌افزاری کاهش می‌یابد. بسیاری از کارکردهای شبکه‌ای که در گذشته به سخت‌افزار نیاز داشتند، اکنون می‌توانند از طریق نرم‌افزار پیاده‌سازی شوند.

با پیشرفت‌های روزافزون در دنیای شبکه‌های کامپیوتری و نیاز به مدیریت بهتر منابع در شبکه‌های پیچیده، مفاهیمی همچون شبکه‌ی تعریف‌شده با نرم‌افزار (SDN) به عنوان یک راه حل نوین مطرح شده‌اند. SDN به عنوان یک معماری نوین، هدفش این است که به شبکه‌ها انعطاف‌پذیری، مقیاس‌پذیری و قابلیت برنامه‌ریزی بدهد، بدون اینکه نیاز به سخت‌افزار پیچیده و پرهزینه باشد. با استفاده از SDN، مدیران شبکه قادرند تمامی عملیات‌های شبکه‌ای را به صورت متمرکز از طریق نرم‌افزار کنترل کنند.

SDN یا شبکه‌ی تعریف‌شده با نرم‌افزار (Software-Defined Networking) یک معماری شبکه‌ای است که در آن کنترل و مدیریت شبکه از طریق نرم‌افزار و بدون نیاز به سخت‌افزار خاص انجام می‌شود. به عبارت دیگر، SDN شبکه را به شکلی انعطاف‌پذیر و قابل برنامه‌ریزی می‌کند، به طوری که مدیران شبکه می‌توانند سیاست‌ها و تنظیمات را از طریق برنامه‌های نرم‌افزاری تغییر دهند.

معماری SDN

معماری SDN معمولاً به سه لایه تقسیم می‌شود:

۱- لایه‌ی کنترل (Control Layer): در این لایه، یک یا چند کنترلر نرم‌افزاری وجود دارند که از طریق API ها، تصمیمات مربوط به مسیریابی و کنترل جریان داده‌ها را اتخاذ می‌کنند. این کنترلرها بر اساس شرایط شبکه، قوانین جدید را اعمال کرده و تصمیمات مربوط به مدیریت ترافیک را به لایه‌های پایین‌تر می‌دهند.

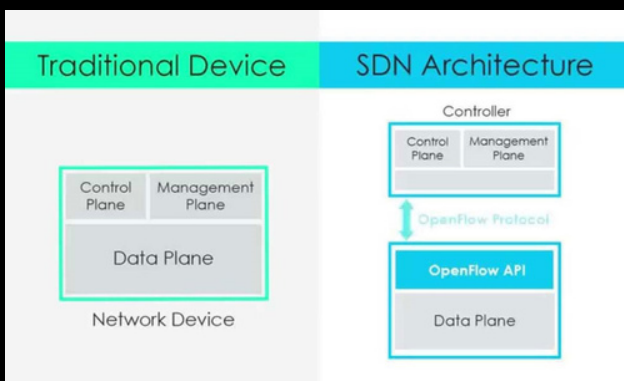
۲- لایه‌ی داده (Data Layer) یا (Forwarding Layer): در این لایه، دستگاه‌های شبکه‌ای «مانند سوئیچ‌ها و روترها» به صورت فیزیکی وجود دارند و بسته‌های داده را طبق دستورات دریافتی از لایه‌ی کنترل پردازش می‌کنند. این لایه وظیفه‌ی ارسال داده‌ها به مقصد مناسب را بر عهده دارد.

در زمان‌های حملات سایبری، می‌توان سیاست‌های امنیتی را به‌صورت فوری اعمال کرد.

۵- Virtualized Network Functions (VNFs): یک SDN با ادغام تکنولوژی‌های مجازی‌سازی مانند NFV (Network Functions Virtualization)، امکان پیاده‌سازی و مدیریت کارکردهای شبکه‌ای مجازی را فراهم می‌آورد، بدون اینکه نیاز به تغییرات سخت‌افزاری باشد. این ویژگی به‌ویژه برای اپراتورهای مخابراتی و مراکز داده مفید است.

۶- موازنه‌ی بار (Load Balancing): SDN می‌تواند به‌طور هوشمند ترافیک شبکه را توزیع کرده و بار را بین سرورها یا دستگاه‌ها به‌طور یکنواخت تقسیم کند. این کار باعث بهبود کارایی و کاهش زمان پاسخ‌دهی در شبکه‌های پر ترافیک می‌شود.

SDN به‌طور کلی معماری جدید و قدرتمندی برای مدیریت شبکه‌ها ارائه می‌دهد که انعطاف‌پذیری، مقیاس‌پذیری، و بهینه‌سازی عملکرد شبکه را افزایش می‌دهد. این تکنولوژی به سازمان‌ها این امکان را می‌دهد که شبکه‌های خود را به‌صورت مرکزی و با استفاده از نرم‌افزار مدیریت کنند و به‌راحتی سیاست‌های شبکه را تغییر دهند. کاربردهای SDN در شبکه‌های بزرگ، IoT، 5G، امنیت و مراکز داده روزبه‌روز بیشتر می‌شود و این تکنولوژی به یکی از ارکان اصلی شبکه‌های مدرن تبدیل شده‌است.



۵- امنیت پیشرفته: SDN می‌تواند ابزارهای پیشرفته‌ای برای نظارت و مدیریت امنیت شبکه فراهم کند، مانند جلوگیری از حملات DDoS، شناسایی تهدیدات و واکنش سریع به حملات. همچنین امکان پیاده‌سازی سیاست‌های امنیتی دقیق‌تر در شبکه به‌کمک نرم‌افزار فراهم است.

۶- بهبود زمان پاسخ‌دهی: به‌دلیل امکان تغییر سریع سیاست‌ها و تنظیمات شبکه، SDN می‌تواند به کاهش زمان تأخیر در واکنش به تغییرات شبکه کمک کند.

کاربردهای SDN

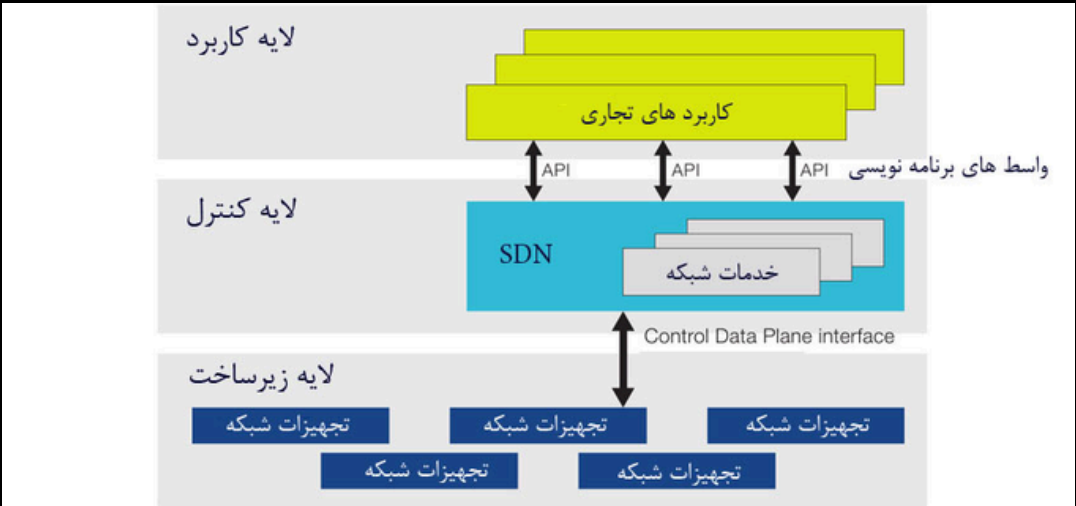
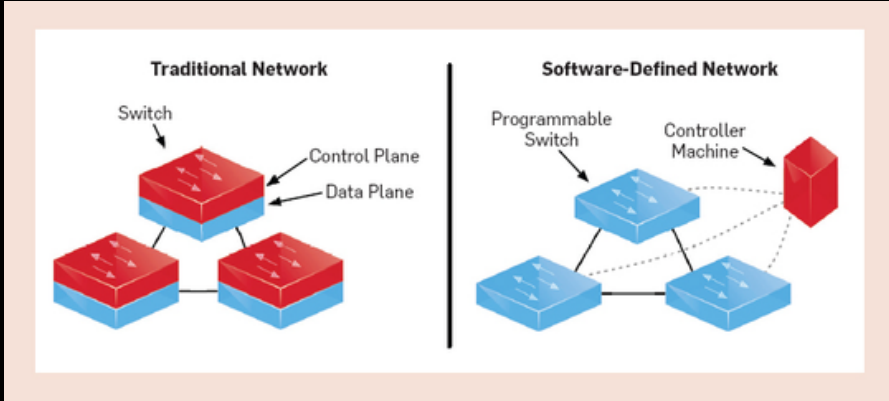
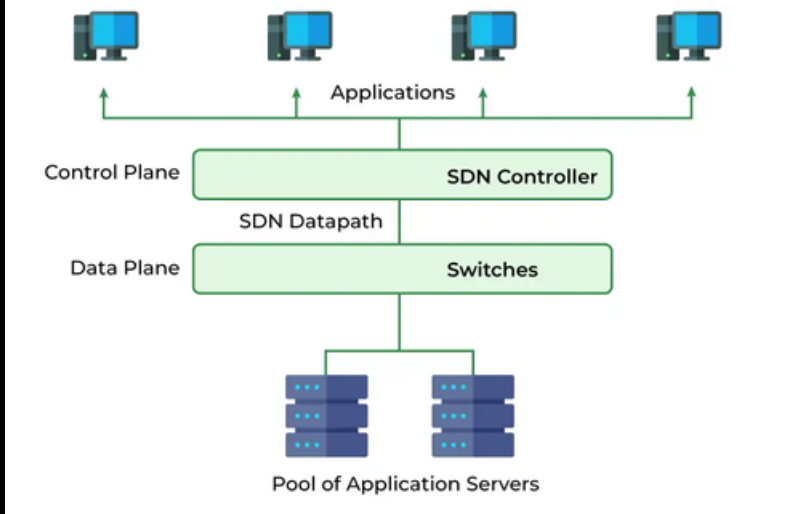
۱- مدیریت شبکه‌های بزرگ و پیچیده: در شبکه‌های بزرگ که شامل هزاران دستگاه و مسیر هستند، SDN می‌تواند کارایی شبکه را با مدیریت متمرکز و شفاف بهبود بخشد. برای مثال، در مراکز داده‌ی بزرگ که نیاز به مدیریت گسترده‌ی ترافیک و تخصیص منابع دارند، SDN امکان مقیاس‌پذیری و بهینه‌سازی را فراهم می‌آورد.

۲- شبکه‌های ابری: در محیط‌های ابری، SDN امکان مدیریت و تخصیص منابع به‌صورت خودکار و با سیاست‌های انعطاف‌پذیر را فراهم می‌کند. این کار باعث بهبود عملکرد و کاهش هزینه‌ها در سرویس‌های ابری می‌شود.

۳- شبکه‌های 5G و IoT: در شبکه‌های 5G، SDN می‌تواند به بهینه‌سازی تخصیص منابع و مدیریت ترافیک کمک کند. همچنین در شبکه‌های اینترنت اشیا (IoT)، که تعداد زیادی دستگاه با ویژگی‌های متفاوت به شبکه متصل هستند، SDN می‌تواند در مدیریت ارتباطات و تخصیص پهنای باند به‌صورت مؤثر عمل کند.

۴- امنیت شبکه: SDN می‌تواند به شناسایی حملات و تهدیدات در شبکه کمک کند و سیاست‌های امنیتی را به‌صورت دینامیک و مطابق با نیازهای شبکه تغییر دهد.

Software Defined Networking (SDN)





۱- تنظیم دستوری (Instruction Tuning): در این روش، مدل‌ها با مجموعه‌ای از دستورات و پاسخ‌های مرتبط آموزش داده می‌شوند. این فرایند به مدل‌ها کمک می‌کند تا توانایی انجام وظایف متنوعی مانند خلاصه‌سازی، ترجمه، یا تحلیل داده‌ها را کسب کنند. به‌طورمثال، یک مدل تنظیم‌شده قادر خواهد بود به سؤالات پیچیده‌ای مانند «چگونه می‌توان یک برنامه‌ی تغذیه‌ی سالم طراحی کرد؟» پاسخ‌های کاربردی دهد.

۲- هم‌ترازی ترجیحات انسانی: در این روش، مدل‌ها برای ارائه‌ی پاسخ‌هایی که با ترجیحات انسانی همخوانی داشته باشند، آموزش داده می‌شوند. این فرایند معمولاً با استفاده از یادگیری تقویتی و بازخورد انسانی (RLHF) انجام می‌شود. به این ترتیب، مدل‌ها می‌توانند پاسخ‌هایی اخلاقی‌تر و مناسب‌تر ارائه دهند. برای مثال، یک مدل پایه ممکن است به سؤال «چگونه می‌توان سیب‌ها را ذخیره کرد؟» پاسخ غیرمرتبط یا ناقصی دهد. اما مدلی که تنظیم دستوری شده باشد، می‌تواند پاسخ‌هایی دقیق‌تر مانند «سیب‌ها باید در محیط خشک و خنک نگهداری شوند» ارائه دهد. دستوردهی (Prompting)، ابزار قدرتمند هدایت مدل‌ها دستوردهی یکی از مؤثرترین ابزارها برای هدایت مدل‌های زبانی است. در این فرایند، کاربر با ارائه‌ی متنی به مدل، آن را برای انجام وظیفه‌ای خاص هدایت می‌کند. این متن که به‌عنوان (پرامپت) شناخته می‌شود، می‌تواند شامل دستورالعمل‌ها، سؤالات، یا مثال‌های کاربردی باشد.

انواع دستوردهی

دستورات ساده: این دستورات شامل جملاتی کوتاه و مستقیم است، مانند: «این متن را به زبان فرانسوی ترجمه کن.» یا «لطفاً یک خلاصه از متن زیر تهیه کن.»

مدل‌های زبانی بزرگ (Large Language models)، تحولی عظیم در پردازش زبان طبیعی ایجاد کرده‌اند. این مدل‌ها اکنون به‌عنوان ابزارهایی هوشمند برای انجام وظایف متنوع از ترجمه‌ی زبان تا تولید متن‌های خلاقانه به کار می‌روند. این مقاله به معرفی مفاهیمی همچون هم‌ترازی مدل، استفاده از دستورات (Prompting) و یادگیری در متن (In-Context Learning) پرداخته و با ارائه‌ی مثال‌ها و تصاویر، کاربردها و چالش‌های مرتبط با این فناوری را بررسی می‌کند.

در دنیای امروز، تکنولوژی‌های هوش مصنوعی به‌سرعت در حال پیشرفت هستند و مدل‌های زبانی بزرگ نقش کلیدی در این پیشرفت‌ها ایفا می‌کنند. این مدل‌ها با پردازش مقادیر عظیمی از داده‌ها، توانایی درک و تولید زبان انسانی را به دست آورده‌اند. با این حال، برای استفاده مؤثر از این مدل‌ها، نیازمند درک ابزارها و تکنیک‌هایی مانند (Prompting) و (In-Context Learning) هستیم که می‌توانند تعامل ما با این فناوری‌ها را بهبود بخشند.

این مقاله تلاش دارد تا این مفاهیم را به زبانی ساده تشریح کرده و با ارائه‌ی مثال‌های واقعی نشان دهد که چگونه می‌توان از این فناوری‌ها در زندگی روزمره و محیط‌های کاری استفاده کرد.

هم‌ترازی مدل: تنظیم برای اهداف انسانی

مدل‌های زبانی بزرگ در اصل برای پیش‌بینی کلمات طراحی شده‌اند. این مدل‌ها با پردازش داده‌های متنی و یادگیری از آن‌ها قادر به تولید متن‌های معنادار هستند؛ اما پیش‌بینی کلمات به‌تنهایی نمی‌تواند تضمین‌کننده‌ی تولید محتوای دقیق و مفید باشد. برای نزدیک کردن خروجی این مدل‌ها به نیازهای انسانی، از تکنیک‌های هم‌ترازی مدل استفاده می‌شود. دو روش کلیدی برای هم‌ترازی مدل‌ها عبارت‌اند از:

دستورات چندمرحله‌ای: در این روش، چند مثال یا توضیح برای کمک به درک بهتر مدل از وظیفه ارائه می‌شود. این تکنیک که به آن (Few-Shot Prompting) نیز گفته می‌شود، مدل را قادر می‌سازد تا عملکرد بهتری داشته باشد.

کاربردهای Prompting

- ۱- خلاصه‌سازی مقالات
- ۲- ترجمه زبان‌ها در مکالمات بین‌المللی
- ۳- تحلیل احساسات کاربران در شبکه‌های اجتماعی
- ۴- تولید محتوای خلاقانه برای نویسندگان و تولیدکنندگان محتوا

یکی از چالش‌های مهم در دستوردهی، طراحی دستورات کارآمد است. دستورات باید به گونه‌ای طراحی شوند که مدل را به سمت تولید خروجی مطلوب هدایت کنند. به‌عنوان مثال، یک دستور مبهم ممکن است منجر به تولید پاسخ‌های نامربوط شود.

یادگیری در متن (In-Context Learning)، یادگیری بدون به‌روزرسانی پارامترها

یکی از ویژگی‌های شگفت‌انگیز مدل‌های زبانی بزرگ، توانایی یادگیری از متن ارائه‌شده در لحظه است. این فرایند که به «یادگیری در متن» معروف است، شامل ارائه مثال‌هایی در قالب متن به مدل برای بهبود عملکرد آن است.

در این روش، کاربر یک یا چند مثال مرتبط با وظیفه‌ی موردنظر را در متن دستور ارائه می‌دهد. مدل با تحلیل این مثال‌ها، رفتار مناسب را برای انجام وظیفه‌ی جدید یاد می‌گیرد. از مزایای این روش می‌توان به انعطاف‌پذیری بالا در انجام وظایف مختلف، عدم نیاز به به‌روزرسانی مدل یا تغییر پارامترها و صرفه‌جویی در زمان و منابع محاسباتی اشاره کرد.

برای مثال، فرض کنید می‌خواهید احساسات یک متن را تحلیل کنید. ابتدا چند نمونه از متن‌های مثبت و منفی به مدل ارائه می‌دهید، سپس مدل بر اساس این مثال‌ها، احساس متن جدید را تحلیل می‌کند.

سرهای القایی (Induction Heads)، مکانیزم‌های کلیدی در یادگیری متن

سرهای القایی یکی از مکانیزم‌های اساسی در مدل‌های زبانی بزرگ هستند که به این مدل‌ها کمک می‌کنند تا الگوهای تکراری را شناسایی و پیش‌بینی کنند. به‌عنوان مثال، اگر مدل دنباله‌ای از کلمات مانند A، B و... ببیند، می‌تواند پیش‌بینی کند که کلمه بعدی (B) خواهد بود.

سرهای القایی چگونه کار می‌کنند؟

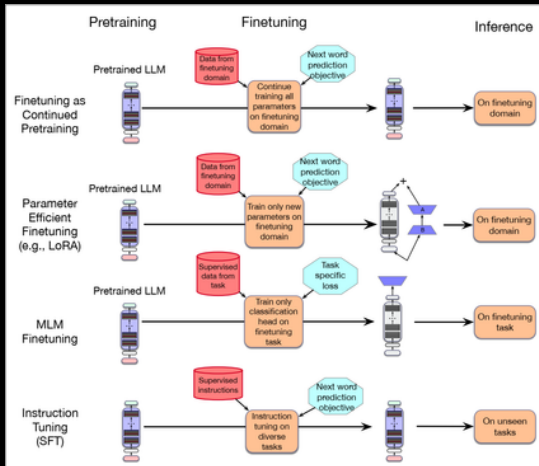
سرهای القایی با استفاده از توجه (Attention) به دنباله‌های قبلی، کلمه‌ای که به‌صورت تکراری در متن ظاهر شده‌است را شناسایی کرده و کلمه‌ی بعدی مرتبط با آن را پیشنهاد می‌دهند.



An induction head looking at vintage uses the prefix matching mechanism to find a prior instance of vintage, and the copying mechanism to predict that cars will occur again.

اهمیت در یادگیری متن

این مکانیزم به مدل کمک می‌کند تا حتی بدون تغییر پارامترها، اطلاعات بیشتری از متن دریافت کرده و عملکرد بهتری داشته باشد.



تنظیم مدل‌ها پس از آموزش اولیه (Post-Training) (and Model Alignment)

مدل‌های زبانی بزرگ پس از مرحله‌ی پیش‌آموزش، نیاز به تنظیمات بیشتری دارند تا بتوانند به نیازهای خاص انسانی پاسخ دهند. این تنظیمات شامل دو مرحله‌ی اصلی است:

- 1- تنظیم دستوری: مدل با استفاده از داده‌های شامل دستورات و پاسخ‌های مرتبط، بهبود می‌یابد.
- 2- هم‌ترازی ترجیحات انسانی: با استفاده از بازخورد انسانی، مدل برای تولید پاسخ‌های ایمن‌تر و مرتبط‌تر تنظیم می‌شود.

Instruction tuning compared to the other kinds of finetuning

دستورالعمل‌ها (Instructions as Training Data) به‌عنوان داده‌های آموزشی

در این مرحله، دستورات به‌عنوان داده‌های آموزشی برای تقویت مدل استفاده می‌شوند. این دستورات می‌توانند شامل توضیحات ساده‌ای مانند «این متن را به زبان چینی ترجمه کن» یا نمونه‌های پیچیده‌تری با محدودیت‌هایی در طول متن یا فرضیه‌های خاص باشند. استفاده از مجموعه‌های عظیم دستورات، شامل هزاران نمونه از وظایف مختلف، به مدل‌ها کمک می‌کند تا نه تنها وظایف متنوعی را بیاموزند، بلکه به شکلی هوشمندانه‌تر به نیازهای کاربران پاسخ دهند.

Chain-of-Thought Prompting:

تصور کنید بخواهید مسئله‌ای پیچیده را حل کنید؛ اولین قدم چیست؟ شکستن مسئله به مراحل کوچک‌تر! این دقیقاً همان ایده‌ای است که در این تکنیک به کار گرفته می‌شود. در این روش، با ارائه‌ی مراحل استدلالی و توضیحی در دستورات آموزشی، مدل تشویق می‌شود که برای حل مسائل، پاسخ‌ها را گام‌به‌گام ارائه دهد. مطالعات نشان داده‌اند که این روش می‌تواند دقت مدل‌ها را در وظایف استدلالی مانند مسائل ریاضی یا استنتاج منطقی به‌طور چشم‌گیری افزایش دهد.

<p>Prompt: Explain the moon landing to a six year old in a few sentences. Output: Explain the theory of gravity to a 6 year old.</p> <hr/> <p>Prompt: Translate to French: The small dog Output: The small dog crossed the road.</p>
--

نمونه‌های عملی و واقعی

- 1- تحلیل احساسات: شناسایی نظرات مثبت یا منفی در نظرسنجی‌ها و بازخوردهای کاربران.
- 2- ترجمه‌ی زبانی: ترجمه‌ی سریع و دقیق متون رسمی و غیررسمی در مکالمات و مستندات.
- 3- خلاصه‌سازی: تهیه‌ی خلاصه‌ای کوتاه از یک گزارش طولانی برای تصمیم‌گیری‌های مدیریتی.
- 4- پاسخ به سؤالات: مدل‌ها می‌توانند به سؤالات تخصصی در حوزه‌هایی مانند پزشکی، حقوق و علوم پاسخ دهند.

تنظیم دستورات (Instruction Tuning)

تنظیم دستورات، یکی از مهم‌ترین روش‌ها برای هم‌راستاسازی مدل‌های زبانی با انتظارات انسانی است. این فرایند شامل آموزش مدل با مجموعه‌ای گسترده از دستورات و پاسخ‌های مرتبط است که به مدل امکان می‌دهد در طیف وسیعی از وظایف، عملکردی مؤثر داشته باشد.

توسعه‌ی دستورات (Prompt Expansion)

توسعه‌ی دستورات، فرایندی است که به منظور تولید نسخه‌های متنوع و بهینه‌تر از یک دستور اولیه انجام می‌شود. این فرایند معمولاً شامل استفاده از مدل‌های زبانی برای ایجاد تغییرات در دستور اصلی است، به‌گونه‌ای که معنای کلی آن حفظ شود. یکی از روش‌های رایج برای این کار، استفاده از (meta prompt) است که مدل زبانی را برای تولید دستورهای جدید هدایت می‌کند. به‌عنوان مثال، از مدل خواسته می‌شود تا با توجه به دستور فعلی، نسخه‌های دیگری ایجاد کند که ممکن است در برخی وظایف عملکرد بهتری داشته باشند. این روش نه تنها به بهبود عملکرد مدل کمک می‌کند، بلکه انعطاف‌پذیری آن را در انجام وظایف مختلف افزایش می‌دهد.

Evaluating Prompted Language Models

ارزیابی مدل‌های زبانی که با دستورات کار می‌کنند (Prompted LMs)، یکی از مراحل مهم برای سنجش دقت و کارایی آن‌ها در انجام وظایف مختلف است. این فرایند معمولاً شامل آزمایش مدل در مجموعه‌ای از وظایف خاص، مانند دسته‌بندی، ترجمه، یا حل مسائل پیچیده است.

یکی از روش‌های رایج برای ارزیابی، استفاده از داده‌های آزمون چندگزینه‌ای مانند MMLU است که در آن، مدل باید به سؤالاتی از حوزه‌های مختلف علمی و عملی پاسخ دهد. نتایج این آزمون‌ها با معیارهایی مانند دقت (Accuracy) یا امتیازات خاص وظایف مانند ROUGE، برای خلاصه‌سازی اندازه‌گیری می‌شود. این ارزیابی‌ها کمک می‌کنند تا نقاط ضعف مدل شناسایی شده و راه‌حل‌هایی برای بهبود عملکرد آن پیشنهاد شوند.

چالش‌های LLMs

۱- امنیت: مدل‌های زبانی ممکن است محتوای مضر تولید کنند؛ بنابراین، لازم است از تکنیک‌هایی مانند هم‌ترازی ترجیحات انسانی برای کاهش خطرات استفاده شود.

Standard Prompting	Chain-of-Thought Prompting
Model Input Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now? A: The answer is 11. Q: The cafeteria had 23 apples. If they used 20 to make lunch and bought 6 more, how many apples do they have?	Model Input Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now? A: Roger started with 5 balls. 2 cans of 3 tennis balls each is 6 tennis balls. $5 + 6 = 11$. The answer is 11. Q: The cafeteria had 23 apples. If they used 20 to make lunch and bought 6 more, how many apples do they have?
Model Output A: The answer is 27. ❌	Model Output A: The cafeteria had 23 apples originally. They used 20 to make lunch. So they had $23 - 20 = 3$. They bought 6 more apples, so they have $3 + 6 = 9$. The answer is 9. ✅

Example of the use of chain-of-thought prompting (right) versus standard prompting (left) on math word problems

بهینه‌سازی خودکار درخواست‌ها (Automatic Prompt Optimization)

فرض کنید می‌خواهید بهترین دستور را برای انجام یک وظیفه بیابید. اما از کجا باید شروع کرد؟ «بهینه‌سازی خودکار دستورات» راهکاری هوشمندانه برای یافتن دستورات مؤثرتر است. در این فرایند، دستورات اولیه (چه توسط انسان و چه به صورت خودکار ایجاد شده باشند) به صورت سیستماتیک تغییر داده می‌شوند و نسخه‌های بهینه‌تری از آن‌ها برای بهبود عملکرد مدل انتخاب می‌شود.

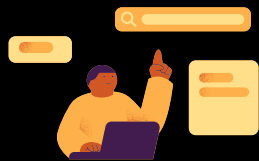
(Candidate Scoring) امتیازدهی به گزینه‌های پیشنهادی

برای انتخاب بهترین دستورات، ابتدا باید آن‌ها را ارزیابی کرد. در این روش، دستورات کاندیدا با استفاده از داده‌های آموزشی آزمایش شده و عملکردشان بر اساس معیارهایی مانند دقت، شباهت به پاسخ‌های صحیح یا معیارهای تخصصی نظیر ROUGE و Bleu سنجیده می‌شود. این ارزیابی سریع و هدفمند به یافتن دستوراتی کمک می‌کند که نه تنها مؤثر باشند، بلکه هزینه‌ی پردازشی کمتری نیز داشته باشند.



نویسندگان:

امیررضا دانشور - دانشجوی کارشناسی ارشد مهندسی کامپیوتر - گرایش هوش مصنوعی
 سعید رفعتی - دانشجوی کارشناسی ارشد مهندسی کامپیوتر - گرایش هوش مصنوعی
 بهنام قهرمان خانی - دانشجوی کارشناسی ارشد مهندسی کامپیوتر - گرایش هوش مصنوعی
 سروش قلی‌زاده - دانشجوی کارشناسی ارشد مهندسی کامپیوتر - گرایش هوش مصنوعی
 امیررضا پورکریم - دانشجوی کارشناسی ارشد مهندسی کامپیوتر - گرایش هوش مصنوعی



منابع:

Jurafsky, D., & Martin, J. H. (2024). *Speech and Language Processing*
<https://B2n.ir/m45579>

۲- دقت: مدل‌ها ممکن است اطلاعات غلط یا گمراه‌کننده تولید کنند. توسعه‌ی مدل‌هایی که بتوانند صحت اطلاعات را تضمین کنند، چالشی مهم است.

۳- حریم خصوصی: استفاده از داده‌های کاربران باید با رعایت قوانین حریم خصوصی انجام شود.

۴- سوگیری: مدل‌ها ممکن است تحت تأثیر داده‌های آموزشی، سوگیری‌هایی داشته باشند که باید شناسایی و برطرف شوند.

مدل‌های زبانی بزرگ ابزارهای قدرتمندی برای حل مسائل پیچیده هستند. استفاده‌ی مؤثر از این فناوری‌ها نیازمند درک دقیق از مفاهیمی مانند دستوردهی و یادگیری در متن است. این فناوری‌ها می‌توانند در حوزه‌های مختلفی مانند آموزش، کسب‌وکار و پژوهش تحولی ایجاد کنند. با این حال، توجه به مسائل اخلاقی و چالش‌های مرتبط با این مدل‌ها ضروری است.

در این مقاله، نقش مدل‌های زبانی بزرگ (LLMs) در تحول پردازش زبان طبیعی و کاربردهای گسترده آن‌ها از ترجمه و تحلیل احساسات تا حل مسائل پیچیده بررسی شد. مفاهیم کلیدی همچون تراز مدل، دستوردهی، و یادگیری در متن نشان دادند که چگونه می‌توان این فناوری‌ها را برای ارائه‌ی خروجی‌های دقیق‌تر و متناسب‌تر با نیازهای انسانی به کار گرفت. با وجود پتانسیل‌های بی‌نظیر این مدل‌ها، چالش‌هایی مانند امنیت، دقت، حریم خصوصی، و سوگیری نیز شناسایی شدند که رفع آن‌ها مستلزم طراحی دقیق، بازخورد انسانی، و استفاده‌ی از تکنیک‌های پیشرفته است. این فناوری، با استفاده‌ی مؤثر و اخلاقی، می‌تواند تحولی بنیادین در حوزه‌های مختلف آموزشی، کسب‌وکار، و پژوهش ایجاد کند.



نمایش‌های پیچیده از داده‌ها، در حوزه‌های مختلفی مانند تشخیص تصویر، پیش‌بینی داده‌ها و به‌ویژه پردازش زبان طبیعی، کاربرد گسترده‌ای پیدا کرده‌اند.

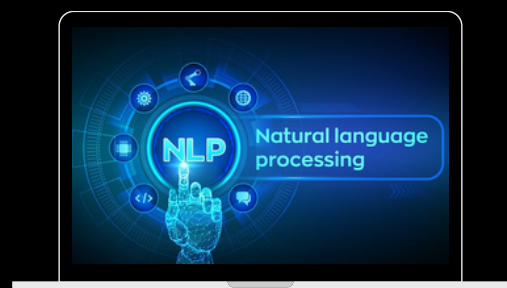
در سال‌های اخیر، پیشرفت‌های قابل‌توجهی در حوزه‌ی NLP با استفاده از شبکه‌های عصبی عمیق (Deep Neural Networks) و مدل‌های مبتنی بر ترانسفورمر (Transformers) مانند BERT و GPT صورت گرفته‌است. این مدل‌ها با استفاده از مکانیزم‌های توجه (Attention Mechanisms) و یادگیری انتقالی (Transfer Learning)، توانسته‌اند عملکرد سیستم‌های NLP را به‌طور چشم‌گیری بهبود بخشند. برای مثال، مدل GPT-3 توانسته است با تولید متن‌های هماهنگ و مرتبط، مرزهای هوش مصنوعی را در حوزه‌ی تولید محتوا گسترش دهد. [1]

با این حال، چالش‌هایی مانند نیاز به داده‌های بزرگ، تفسیرپذیری مدل‌ها و مسائل اخلاقی همچنان وجود دارند. این مقاله به بررسی کاربردهای شبکه‌های عصبی در NLP و چالش‌های پیش روی این حوزه می‌پردازد و زمینه‌های کاری آینده را نیز مورد بررسی قرار می‌دهد.

معرفی شبکه‌های عصبی

شبکه‌های عصبی مصنوعی از واحدهایی به نام نورون تشکیل شده‌اند که هر نورون شامل ورودی‌ها، وزن‌ها، یک تابع جمع وزن‌دار و یک تابع فعال‌سازی غیرخطی است. این شبکه‌ها با ترکیب لایه‌های متعدد، قادر به یادگیری نمایش‌های پیچیده از داده‌ها هستند. در حوزه‌ی NLP، شبکه‌های عصبی پیش‌خور (Feedforward Neural Networks) و شبکه‌های عمیق (Deep Neural Networks) به‌طور گسترده‌ای مورد استفاده قرار می‌گیرند. توابع فعال‌سازی غیرخطی مانند ReLU و سیگموئید، نقش کلیدی در ایجاد خروجی‌های غیرخطی و بهبود عملکرد شبکه‌های عصبی ایفا می‌کنند.

شبکه‌های عصبی مصنوعی (ANNs) به‌عنوان یکی از ابزارهای قدرتمند در حوزه‌ی یادگیری ماشین و هوش مصنوعی، نقش مهمی در پیشرفت پردازش زبان طبیعی (NLP) ایفا کرده‌اند. این شبکه‌ها با الهام از ساختار مغز انسان، قادر به یادگیری الگوهای پیچیده در داده‌های متنی هستند و در سال‌های اخیر، پیشرفت‌های چشم‌گیری در این حوزه ایجاد کرده‌اند. کاربردهای شبکه‌های عصبی در NLP شامل تحلیل احساسات، مدل‌سازی زبان، ترجمه ماشینی، تولید متن، کلاس‌بندی متن، پیش‌بینی لغات و کاربردهای نوظهور مانند پاسخ‌گویی به سؤالات و خلاصه‌سازی متن است. این مقاله به بررسی این کاربردها می‌پردازد و چالش‌های پیش روی این حوزه را نیز مورد بررسی قرار می‌دهد. درنهایت، زمینه‌های کاری آینده و مسائل باز در این حوزه ارائه خواهد شد.



پردازش زبان طبیعی (NLP) یکی از شاخه‌های مهم هوش مصنوعی است که به تعامل بین کامپیوترها و زبان انسان‌ها می‌پردازد. با گسترش داده‌های متنی و نیاز به تحلیل خودکار آن‌ها، شبکه‌های عصبی مصنوعی به‌عنوان ابزاری کارآمد برای حل مسائل پیچیده‌ی NLP مورد استفاده قرار گرفته‌اند. این شبکه‌ها با الهام از ساختار مغز انسان، قادر به یادگیری الگوهای پیچیده در داده‌های متنی هستند و در سال‌های اخیر، پیشرفت‌های چشم‌گیری در این حوزه ایجاد کرده‌اند. شبکه‌های عصبی به‌دلیل توانایی در یادگیری



کاربرد شبکه‌های عصبی در پردازش زبان طبیعی

شبکه‌های عصبی در حوزه‌ی NLP برای حل مسائل مختلفی مانند تحلیل احساسات، ترجمه‌ی ماشینی، تولید متن، مدل‌سازی زبان، کلاس‌بندی متن، پیش‌بینی لغات و کاربردهای نوظهور استفاده می‌شوند. در ادامه، این کاربردها به تفصیل شرح داده می‌شوند.

۱- تحلیل احساسات (Sentiment Analysis): تحلیل احساسات یکی از کاربردهای مهم شبکه‌های عصبی در NLP است. این فرایند شامل تشخیص احساسات مثبت یا منفی موجود در یک متن است. شبکه‌های عصبی با استفاده از ویژگی‌های متنی مانند کلمات و عبارات، قادر به پیش‌بینی احساسات هستند. برای مثال، در تحلیل نظرات کاربران درباره‌ی یک محصول، شبکه‌های عصبی می‌توانند تشخیص دهند که آیا نظر کاربر مثبت است یا منفی.

یکی از مزایای استفاده از شبکه‌های عصبی در تحلیل احساسات، توانایی آن‌ها در یادگیری ویژگی‌های پیچیده و غیرخطی از داده‌ها است. به جای استفاده از ویژگی‌های دست‌ساز، شبکه‌های عصبی می‌توانند به‌طور خودکار ویژگی‌های مفید را از متن استخراج کنند. این ویژگی‌ها می‌توانند شامل کلمات کلیدی، عبارات و حتی ساختارهای نحوی باشند. برای مثال، در تحلیل احساسات، شبکه‌های عصبی می‌توانند تشخیص دهند که کلماتی مانند «عالی» یا «بد» نشان‌دهنده‌ی احساسات مثبت یا منفی هستند. [1]

۲- مدل‌سازی زبان (Language Modeling): مدل‌سازی زبان یکی دیگر از کاربردهای مهم شبکه‌های عصبی در NLP است. این فرایند شامل پیش‌بینی کلمه‌ی بعدی در یک دنباله‌ی متنی بر اساس کلمات قبلی است. مدل‌های زبانی عصبی، برخلاف مدل‌های سنتی مانند N-gram، قادر به درک بهتر کونکست متن هستند و می‌توانند حتی در جملات پیچیده و طولانی نیز عملکرد خوبی داشته باشند. برای مثال، مدل GPT-3 می‌تواند با دقت بالایی کلمات بعدی را در یک جمله پیش‌بینی کند و حتی متن‌های طولانی و پیچیده را تولید کند. [1,2]

۳- ترجمه‌ی ماشینی (Machine Translation): ترجمه‌ی ماشینی یکی از قدیمی‌ترین و مهم‌ترین کاربردهای شبکه‌های عصبی در NLP است. شبکه‌های عصبی، به‌ویژه مدل‌های مبتنی بر ترانسفورمر (Transformers)، توانسته‌اند کیفیت ترجمه‌ی ماشینی را به‌طور چشم‌گیری بهبود بخشند. این مدل‌ها با استفاده از مکانیزم‌های توجه (Attention Mechanisms)، قادر به درک بهتر ساختار جمله و ترجمه‌ی دقیق‌تر آن به زبان مقصد هستند. برای مثال، مدل Transformer گوگل، ترجمه‌های بسیار دقیق و روانی را ارائه می‌دهد. [3]

۴- تولید متن (Text Generation): تولید متن یکی از کاربردهای جذاب شبکه‌های عصبی در NLP است. این فرایند شامل تولید متن‌های جدید بر اساس الگوهای یادگرفته‌شده از داده‌های متنی است. شبکه‌های عصبی، به‌ویژه مدل‌های مبتنی بر GPT، قادر به تولید متن‌های هماهنگ و مرتبط هستند. این کاربرد در سیستم‌های تولید محتوای خودکار، ربات‌های چت و حتی تولید شعر و داستان استفاده می‌شود. [4]

۵- کلاس‌بندی متن (Text Classification): کلاس‌بندی متن یکی از کاربردهای اصلی شبکه‌های عصبی در NLP است. این فرایند شامل اختصاص

برچسب‌های از پیش تعریف‌شده به متون بر اساس محتوای آن‌ها است. برای مثال، در تحلیل احساسات، متن‌ها به دسته‌های مثبت، منفی یا خنثی تقسیم می‌شوند. شبکه‌های عصبی، به‌ویژه مدل‌های عمیق مانند شبکه‌های عصبی پیچشی (CNNS) و شبکه‌های عصبی بازگشتی (RNNS)، به دلیل توانایی در یادگیری ویژگی‌های پیچیده و غیرخطی از داده‌های متنی، در این زمینه بسیار مؤثر هستند. [5] برچسب‌های از پیش تعریف‌شده به متون بر اساس محتوای آن‌ها است. برای مثال، در تحلیل احساسات، متن‌ها به دسته‌های مثبت، منفی یا خنثی تقسیم می‌شوند. شبکه‌های عصبی، به‌ویژه مدل‌های عمیق مانند شبکه‌های عصبی پیچشی (CNNS) و شبکه‌های عصبی بازگشتی (RNNS)، به دلیل توانایی در یادگیری ویژگی‌های پیچیده و غیرخطی از داده‌های متنی، در این زمینه بسیار مؤثر هستند. [5]

۶- پیش‌بینی لغات (Word Prediction): پیش‌بینی لغات یکی دیگر از کاربردهای مهم شبکه‌های عصبی در پردازش زبان طبیعی است. این فرایند شامل پیش‌بینی کلمه‌ی بعدی در یک دنباله‌ی متنی بر اساس کلمات قبلی است. این کاربرد در سیستم‌های تکمیل خودکار متن، ترجمه‌ی ماشینی و دستیارهای صوتی مانند الکسا و گوگل اسیستنت بسیار مفید است. شبکه‌های عصبی بازگشتی (RNNS) و مدل‌های مبتنی بر ترانسفورمر (Transformers) مانند GPT در پیش‌بینی لغات بسیار مؤثر هستند. [7]

۷- کاربردهای نوظهور (Emerging Applications): علاوه بر کاربردهای سنتی، شبکه‌های عصبی در حوزه‌های نوظهور NLP نیز نقش مهمی ایفا می‌کنند. برای مثال، در سیستم‌های پاسخ‌گویی به سؤالات (Question Answering)، شبکه‌های عصبی قادر به درک سؤالات پیچیده و ارائه‌ی پاسخ‌های دقیق هستند.

همچنین، در حوزه‌هایی مانند خلاصه‌سازی متن (Text Summarization) و تشخیص موجودیت‌های نام‌گذاری‌شده (Named Entity Recognition)، شبکه‌های عصبی عملکرد قابل‌توجهی از خود نشان داده‌اند. [8]

شبکه‌های عصبی در پردازش زبان طبیعی نقش بسیار مهمی ایفا می‌کنند. کاربردهای این شبکه‌ها، از تحلیل احساسات و مدل‌سازی زبان تا ترجمه‌ی ماشینی و تولید متن، نشان‌دهنده‌ی توانایی آن‌ها در یادگیری الگوهای پیچیده از داده‌های متنی است. با پیشرفت‌های اخیر در فناوری‌های شبکه‌های عصبی عمیق، انتظار می‌رود که این شبکه‌ها در آینده نقش مهم‌تری در حوزه‌ی NLP ایفا کنند. به‌ویژه، استفاده از توابع فعال‌سازی غیرخطی مانند ReLU و بهبود مدل‌های زبانی عصبی، می‌تواند به پیشرفت‌های بیشتری در این حوزه منجر شود.

زمینه‌های کاری در آینده

با وجود پیشرفت‌های چشم‌گیر در حوزه‌ی NLP، هنوز چالش‌ها و مسائل باز زیادی وجود دارند که نیاز به تحقیقات بیشتر دارند. برخی از این چالش‌ها عبارت‌اند از: ۱- تفسیرپذیری مدل‌ها: یکی از چالش‌های اصلی در شبکه‌های عصبی، تفسیرپذیری مدل‌ها است. در حالی که این مدل‌ها عملکرد خوبی دارند، درک نحوه‌ی تصمیم‌گیری آن‌ها هنوز دشوار است. تحقیقات آینده باید بر روی توسعه‌ی مدل‌های قابل‌تفسیرتر متمرکز شود. [7,8]

۲- نیاز به داده‌های کمتر: آموزش مدل‌های عصبی به حجم زیادی از داده‌ها نیاز دارد. توسعه‌ی روش‌هایی که بتوانند با داده‌های کمتر نیز عملکرد خوبی داشته باشند، یکی از زمینه‌های تحقیقاتی مهم است. [9]

۳- مسائل اخلاقی: با افزایش استفاده از شبکه‌های عصبی در NLP، مسائل اخلاقی مانند سوگیری (Bias) در مدل‌ها و حریم خصوصی داده‌ها نیز اهمیت بیشتری

[5] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.



[6] Goldberg, Y. (2017). Neural Network Methods for Natural Language Processing. Morgan & Claypool Publishers.

نویسندگان:

[7] Liu, Y., Ott, M., Goyal, N., et al. (2019). RoBERTa: A Robustly Optimized BERT Pretraining Approach. arXiv preprint arXiv:1907.11692.

[8] Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why Should I Trust You?": Explaining the Predictions of Any Classifier. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.

[9] Wang, A., Singh, A., Michael, J., et al. (2018). GLUE: A Multi-Task Benchmark and Analysis Platform for Natural Language Understanding. arXiv preprint arXiv:1804.07461.

[10] Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?. Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency.

[11] Baltrušaitis, T., Ahuja, C., & Morency, L. P. (2019). Multimodal Machine Learning: A Survey and Taxonomy. IEEE Transactions on Pattern Analysis and Machine Intelligence, 41(2), 423-443.

علی اشرف - دانشجوی کارشناسی ارشد مهندسی کامپیوتر - گرایش هوش مصنوعی - ترم دوم
مجتبی پورطهماسبی - دانشجوی کارشناسی ارشد مهندسی کامپیوتر - گرایش هوش مصنوعی - ترم دوم

دانیال غفاری - دانشجوی کارشناسی ارشد مهندسی کامپیوتر - گرایش هوش مصنوعی - ترم دوم
سینا خیراندیش - دانشجوی کارشناسی ارشد مهندسی کامپیوتر - گرایش هوش مصنوعی - ترم دوم

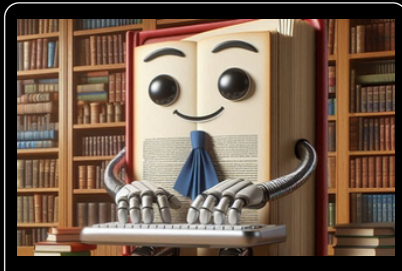


[1] Devlin, J., Chang, M., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv preprint arXiv:1810.04805.

[2] Brown, T., Mann, B., Ryder, N., et al. (2020). Language Models are Few-Shot Learners. Advances in Neural Information Processing Systems, 33, 1877-1901.

[3] Vaswani, A., Shazeer, N., Parmar, N., et al. (2017). Attention is All You Need. Advances in Neural Information Processing Systems, 30, 5998-6008.

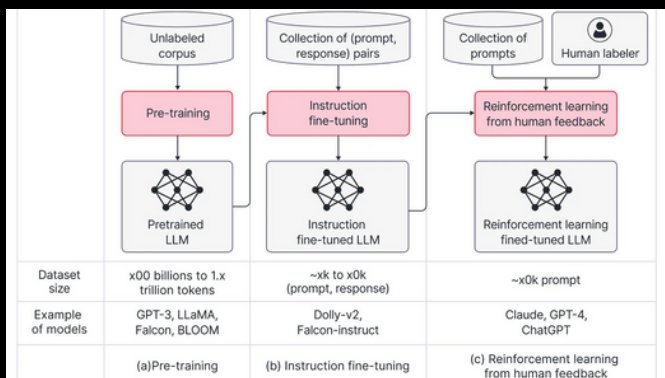
[4] Zhang, L., Wang, S., & Liu, B. (2018). Deep Learning for Sentiment Analysis: A Survey. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 8(4), e1253.



آموزش مدل‌های زبانی بزرگ (LLM ها) به مجموعه‌ی داده‌های گسترده‌ای نیاز دارد که از منابع متنوعی مانند وب، کتاب‌ها و مقالات جمع‌آوری می‌شوند. این داده‌ها پیش از استفاده، برای حذف محتوای نامناسب و ارتقای کیفیت پردازش و فیلتر می‌شوند. همچنین، به منظور بهبود عملکرد LLM ها در وظایف خاص، می‌توان آن‌ها را با استفاده از داده‌های مرتبط با آن وظایف به صورت دقیق تنظیم (fine-tuning) کرد.

آموزش LLM ها شامل سه فاز اصلی می‌باشد:

- ۱- پیش‌آموزش (Pre-training)
 - ۲- تنظیم دقیق (Fine-tuning) یا (Instruction tuning)
 - ۳- یادگیری تقویتی با فیدبک انسانی یا (Reinforcement learning with human feedback)
- این تصویر از سایت OpenAI خیلی واضح‌تر و شفاف‌تر مراحل آموزش را توضیح می‌دهد:



در عصر تحول‌آفرین هوش مصنوعی، یکی از جذاب‌ترین شاخه‌ها، پردازش زبان طبیعی (NLP) است؛ جایی که ماشین‌ها در تلاش‌اند زبان انسان را درک کنند و با آن تعامل داشته باشند. در این میان، مدل‌های زبانی بزرگ (LLM ها) به‌عنوان انقلابی‌ترین پیشرفت در این حوزه شناخته می‌شوند.

LLM ها سیستم‌های هوشمندی هستند که بر پایه‌ی حجم عظیمی از داده‌های متنی آموزش دیده‌اند. این مدل‌ها توانایی انجام وظایف متنوعی مانند ترجمه‌ی ماشینی، خلاصه‌سازی متون، پاسخ‌گویی به سؤالات و تولید متن‌های خلاقانه را دارند. معماری پیشرفته‌ی ترانسفورماتور، قلب تپنده‌ی این مدل‌هاست که آن‌ها را قادر می‌سازد روابط پیچیده‌ی میان کلمات را تحلیل کنند و خروجی‌هایی منسجم و با کیفیت بالا ارائه دهند.

با ظهور این مدل‌ها، دنیای NLP وارد مرحله‌ای جدید از نوآوری شده‌است؛ جایی که ارتباط میان انسان و ماشین عمیق‌تر از همیشه شده و آینده‌ی پر از امکانات تازه را نوید می‌دهد.

یکی از ویژگی‌های برجسته‌ی مدل‌های زبانی بزرگ (LLM ها)، توانایی تولید متن است. برای این منظور، از روش‌های مختلف رمزگشایی (decoding) استفاده می‌شود که از جمله آن‌ها می‌توان به رمزگشایی حریصانه (greedy decoding) و روش‌های نمونه‌برداری (sampling) اشاره کرد. روش‌های نمونه‌برداری مانند نمونه‌برداری تصادفی، نمونه‌برداری top-k و نمونه‌برداری top-p به تولید متون متنوع‌تر و جذاب‌تر کمک می‌کنند.

معرفی مدل‌های زبانی بزرگ (LLM)

۱- پیش‌زمینه: مدل‌های زبانی بزرگ (LLM) ابزارهای پیشرفته‌ای هستند که با بهره‌گیری از حجم عظیمی از داده‌های متنی آموزش دیده‌اند. این مدل‌ها قادرند وظایف پیچیده‌ی پردازش زبان طبیعی (NLP) مانند ترجمه‌ی ماشینی، خلاصه‌سازی متون، پاسخ به سؤالات، و تولید محتوای خلاقانه را با دقت و کیفیت بالا انجام دهند.

۲- معماری: این مدل‌ها معمولاً بر پایه‌ی معماری ترانسفورماتور طراحی شده‌اند؛ ساختاری که توانایی تجزیه و تحلیل روابط پیچیده بین کلمات را دارد و متن‌هایی منسجم و معنادار تولید می‌کند.

۳- فرایند آموزش: آموزش LLM ها با روش خودنظارتی (Self-Supervised Learning) انجام می‌شود. در این روش، مدل‌ها از داده‌های بدون برچسب برای یادگیری و بهبود عملکرد خود بهره می‌گیرند. این فرایند به آن‌ها امکان می‌دهد تا از داده‌های گسترده برای درک و پردازش بهتر زبان استفاده کنند.

۴- تولید متن با LLM ها:

الف) رمزگشایی: برای تولید متن، از تکنیک‌های مختلف رمزگشایی مانند رمزگشایی حریصانه (Greedy Decoding) و نمونه‌برداری (Sampling) استفاده می‌شود.

ب) نمونه‌برداری: روش‌های نمونه‌برداری مانند نمونه‌برداری تصادفی، نمونه‌برداری Top-K و Top-P به تولید متن‌های متنوع و جذاب کمک می‌کنند.

ارزیابی LLM ها با استفاده از معیارهای مختلفی از جمله perplexity و دقت در وظایف پایین‌دستی انجام می‌شود. علاوه بر دقت، عواملی مانند اندازه‌ی مدل، زمان آموزش و استنتاج، مصرف انرژی و ملاحظات عدالت نیز در ارزیابی این مدل‌ها مورد توجه قرار می‌گیرند.

باتوجه به اندازه‌ی بزرگ LLM ها، برای افزایش کارایی در زمان استنتاج، از KV Cache جهت ذخیره‌سازی بردارهای کلید و مقدار استفاده می‌شود که باعث کاهش محاسبات تکراری می‌گردد. همچنین، برای کاهش هزینه‌های محاسباتی در فرایند تنظیم دقیق این مدل‌ها، از روش‌های تنظیم دقیق کارآمد پارامتر مانند LORA بهره گرفته می‌شود.

مدل‌های زبانی بزرگ (LLM ها)، باوجود مزایای بی‌شمار، خطراتی نیز به همراه دارند که نباید نادیده گرفته شوند. این خطرات شامل تولید اطلاعات نادرست یا توهم‌آمیز، به‌کارگیری زبان توهین‌آمیز، نشت اطلاعات خصوصی و سوءاستفاده توسط افراد مخرب است. برای بهره‌گیری مسئولانه از این فناوری پیشرفته، شناخت این چالش‌ها و تلاش برای کاهش آن‌ها ضروری است.

در این راستا، اقدامات مؤثری می‌توان انجام داد: تحلیل دقیق داده‌های مورد استفاده برای آموزش، شفاف‌سازی منابع داده و ترویج استفاده از مدل‌های منبع‌باز. این راهکارها نه تنها خطرات را کاهش می‌دهند، بلکه اعتماد به این فناوری را نیز افزایش می‌دهند.

در نهایت، LLM ها به‌عنوان یکی از مهم‌ترین نوآوری‌های اخیر در پردازش زبان طبیعی (NLP)، پتانسیل ایجاد تحول در حوزه‌های گوناگون را دارند. با درک صحیح مزایا، شناسایی خطرات و اجرای روش‌های کاهش آن‌ها، می‌توانیم از این فناوری قدرتمند به شکلی مسئولانه و مؤثر استفاده کنیم.

۵- آموزش LLM ها:

الف) داده‌های آموزشی: LLM ها بر روی مجموعه‌ی داده‌های عظیم شامل متون جمع‌آوری شده از وب، کتاب‌ها و مقالات آموزش داده می‌شوند. فیلتر کردن داده‌ها: برای بهبود کیفیت و حذف محتوای نامناسب، داده‌ها پیش از آموزش فیلتر می‌شوند. تنظیم دقیق: برای بهینه‌سازی عملکرد در وظایف خاص، تنظیم دقیق (Fine-Tuning) با استفاده از داده‌های مرتبط انجام می‌شود.

۶- ارزیابی عملکرد LLM ها:

الف) معیارهای ارزیابی: مدل‌ها با معیارهایی مانند پیچیدگی (Perplexity) و دقت در وظایف خاص ارزیابی می‌شوند. ب) عوامل دیگر: علاوه بر دقت، عواملی مانند اندازه‌ی مدل، زمان آموزش، مصرف انرژی و عدالت نیز در ارزیابی لحاظ می‌شوند.

۷- مقیاس‌پذیری و بهینه‌سازی:

الف) قوانین مقیاس‌پذیری: عملکرد LLM ها با اندازه‌ی مدل، حجم داده‌های آموزشی و منابع محاسباتی رابطه‌ی مستقیم دارد. ب) افزایش کارایی: استفاده از تکنیک‌هایی مانند KV Cache برای ذخیره‌سازی بردارهای کلید و مقدار، باعث افزایش سرعت پردازش می‌شود. ج) تنظیم دقیق کارآمد پارامترها: روش‌هایی مانند LORA برای کاهش هزینه‌های محاسباتی و بهینه‌سازی عملکرد به کار می‌روند.

۸- خطرات احتمالی LLM ها:

الف) توهم اطلاعات: مدل‌ها ممکن است اطلاعات نادرست یا گمراه‌کننده تولید کنند.

ب) زبان توهین‌آمیز: احتمال تولید محتوای توهین‌آمیز یا کلیشه‌ای وجود دارد.

ج) نشت اطلاعات: اطلاعات خصوصی ممکن است از داده‌های آموزشی نشت پیدا کند.

د) سوءاستفاده: مدل‌ها می‌توانند برای تولید اطلاعات نادرست یا فعالیت‌های مخرب مانند فیشینگ به کار روند.

۹- کاهش خطرات:

الف) تحلیل داده‌ها: داده‌های آموزشی باید برای شناسایی مسائل مرتبط با سمیت، تعصب، و حریم خصوصی بررسی شوند.

ب) شفافیت: ارائه‌ی اطلاعات دقیق درباره داده‌های آموزشی از طریق برگه‌های داده یا کارت‌های مدل ضروری است.

ج) مدل‌های منبع‌باز: استفاده از مدل‌های منبع‌باز، شفافیت بیشتری در داده‌ها و عملکرد مدل ایجاد می‌کند.

جمع‌بندی نهایی

مدل‌های زبانی بزرگ (LLM) ابزارهایی قدرتمند هستند که تحولی شگرف در پردازش زبان طبیعی ایجاد کرده‌اند. با این حال، بهره‌برداری مؤثر و مسئولانه از آن‌ها نیازمند درک عمیق از خطرات احتمالی و تلاش برای کاهش این چالش‌ها است. در آینده، توسعه‌ی این مدل‌ها با تمرکز بر شفافیت، دقت، و امنیت می‌تواند راه را برای پیشرفت‌های بیشتر هموار کند.



نویسندگان:

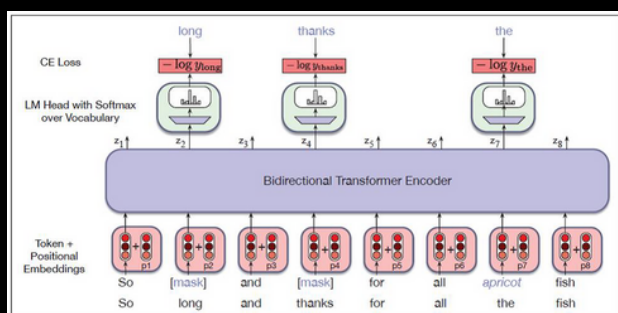
کمیل آقابابایی - دانشجوی کارشناسی ارشد مهندسی کامپیوتر - گرایش هوش مصنوعی - ترم دوم
محمد طالبی - دانشجوی کارشناسی ارشد مهندسی کامپیوتر - گرایش هوش مصنوعی - ترم دوم
محمد علیزاده - دانشجوی کارشناسی ارشد مهندسی کامپیوتر - گرایش هوش مصنوعی - ترم دوم
سروش سهرابی - دانشجوی کارشناسی ارشد مهندسی کامپیوتر - گرایش هوش مصنوعی - ترم دوم
امیر اسدنیا - دانشجوی کارشناسی ارشد مهندسی کامپیوتر - گرایش هوش مصنوعی - ترم دوم



Speech and Language Processing

<https://arxiv.org/abs/2307.06435>

https://www.researchgate.net/publication/372258530_Large_Language_Models_A_Comprehensive_Survey_of_its_Applications_Challenges_Limitations_and_Future_Prospects



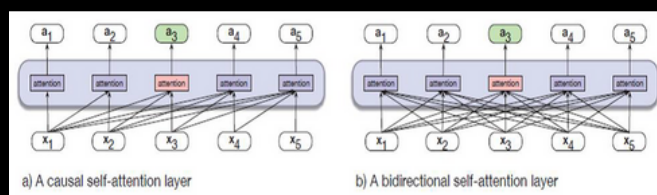
مدل‌های زبانی یک‌طرفه و دوطرفه

مدل‌های زبانی یک‌طرفه، مولد هستند «یعنی برای تولید متن طراحی شده‌اند» و هدف آن‌ها، تولید توکن بعدی در یک توالی است. به این مدل‌ها، Decoder-Only نیز گفته می‌شود. این مدل‌ها نمی‌توانند از اطلاعات موجود توکن‌های آینده، استفاده کنند. مدل‌های زبانی دوطرفه، مولد نیستند بلکه برای محاسبه یا بازنمایی متنی طراحی شده‌اند. به این مدل‌ها، Encoder-Only نیز گفته می‌شود. این مدل‌ها، برای وظایف تفسیری مانند طبقه‌بندی متن، استخراج رابطه‌ها و... مناسب هستند. مدل‌های زبانی دوطرفه، دارای توانایی درک متن در هر دو جهت و بازنمایی‌هایی پیشرفته هستند و نتایج خوبی را ارائه داده‌اند. این مدل‌ها، از مکانیزم Self-Attention استفاده می‌کنند. این مکانیزم، امکان این را فراهم می‌آورد که مدل بر تمام کلمات یک توالی ورودی مروری داشته باشد و بتواند روابط معنایی آن‌ها را درک کند. در این مدل‌ها، هر توکن می‌تواند به تمام توکن‌های دیگر در متن، از جمله توکن‌های قبلی و بعدی توجه کند. این امر سبب می‌شود مدل بتواند اطلاعات متن را به صورت جامع و دقیق‌تری درک کند.

مدل‌های زبانی سنتی و قدیمی‌تر برای انجام یک سری وظایف خاص و درک معنای دقیق کلمات در یک زمینه‌ی خاص مناسب نبودند؛ از این رو وجود مدل‌های زبانی پیشرفته‌تری که بتوانند این چالش را برطرف سازند و توجه به کل ساختار متن داشته باشند، امری ضروری تلقی شد و این امر، سبب پیدایش مدل‌های زبانی نقاب‌دار شد. در این مقاله، به ارائه‌ی مطالبی در خصوص مدل‌های زبانی نقاب‌دار، تفاوت میان مدل‌های زبانی یک‌طرفه و دوطرفه، تعبیه متنی، تکنیک‌های نقاب‌دار کردن و کاربرد مدل‌های زبانی نقاب‌دار، خواهیم پرداخت.

معرفی مدل‌های زبانی نقاب‌دار

مدل‌های زبانی نقاب‌دار یا به اختصار (MLMs) نوعی از مدل‌های یادگیری عمیق هستند که برای درک و تولید زبان طبیعی طراحی شده‌اند. این مدل‌ها معمولاً در پردازش زبان طبیعی استفاده می‌شوند. با استفاده از مدل‌های زبانی نقاب‌دار، معنای کلمات در متن‌های مختلف به طور عمیق درک می‌شود. در این مدل‌ها، بخشی از کلمات موجود در یک جمله با یک توکن خاص (معمولاً [MASK]) جایگزین می‌شوند و مدل تلاش می‌کند کلمات پوشانده‌شده را بر اساس بقیه‌ی کلمات جمله پیش‌بینی کند. استفاده از این روش سبب درک عمیق‌تر مدل از ساختار و معنای زبان می‌شود. در مدل‌های زبانی نقاب‌دار، توجه مدل به کل جمله «هم به سمت چپ و هم به سمت راست توکن‌های نقاب‌دار» متمرکز می‌شود تا بتواند به صورت دوسویه روابط معنایی بین کلمات را یاد بگیرد.



تعبیه متنی مدل‌های زبانی یک‌طرفه و دوطرفه

منظور از تعبیه، نمایش کلمات به شکل یک بردار است که توسط مدل‌های زبانی ساخته می‌شود. در روش‌های سنتی مانند Word2vec یا GloVe به هر کلمه بدون توجه به اینکه در چه متن یا جمله‌ای به کار رفته، یک بردار ثابت اختصاص داده می‌شود؛ اما در مدل‌های زبانی نقاب‌دار مانند BERT، هر کلمه باتوجه‌به متنی که در آن ظاهر می‌شود، یک بردار متفاوت دریافت می‌کند. به‌عنوان مثال، یک کلمه می‌تواند چندین معنی داشته باشد، در چنین حالتی، باتوجه‌به ساختار مدل‌های زبانی نقاب‌دار، کلمات باتوجه‌به جمله‌ای که در آن به کار می‌روند، بردار متفاوتی خواهند داشت. به‌طورکلی، مدل‌های زبانی دوطرفه، در بازنمایی معنای دقیق کلمات، عملکرد بهتری دارند؛ بنابراین استفاده از روش‌های جدیدتر و مدل‌های زبانی نقاب‌دار، این امکان را فراهم ساخته‌است که مدل‌ها بتوانند معنای دقیق کلمات را باتوجه‌به زمینه‌شان بهتر درک کنند؛ اما مدل‌های زبانی سنتی و یک‌طرفه، فقط نمای ثابتی از کلمه را ارائه می‌کردند.

تکنیک‌های نقاب‌دار کردن

نحوهی کار مدل‌های زبانی نقاب‌دار، بر اساس Cloze Task می‌باشد. بدین ترتیب، مدل کلمات حذف‌شده «ماسک‌شده» در یک جمله را پیش‌بینی می‌کند. در فرایند آموزش بر اساس این ساختار، ورودی مدل دستکاری می‌شود و بدین ترتیب وظیفه‌ی مدل، بازسازی متن اصلی از نسخه‌ی دستکاری‌شده می‌باشد. درحقیقت، مدل یاد می‌گیرد که چگونه ورودی اصلی و بدون نویز را بازسازی کند. این روش مدل را مجبور می‌کند تا به کل متن توجه کند و بتواند روابط معنایی بین کلمات را درک کند.

تکنیک‌های نقاب‌دار کردن

۱- جایگذاری (Substitution): در این تکنیک، کلمات صحیح با کلمات اشتباه یا بی‌ربط جایگزین می‌شوند.

۲- حذف کردن (Deletion): در این تکنیک، برخی از کلمات متن حذف می‌شوند.

۳- تغییر ترتیب کلمات (Reordering): در این تکنیک، ترتیب کلمات در جمله تغییر می‌کند.

۴- افزودن کلمات اضافه (Insertion): در این تکنیک، کلمات بی‌ربط به متن اضافه می‌شوند.

کاربرد مدل‌های زبانی نقاب‌دار

۱- طبقه‌بندی متن (Text Classification): بدین صورت، مدل‌های زبانی می‌توانند متون را به دسته‌های مختلف طبقه‌بندی کنند. به‌عنوان مثال، می‌توانند نظرات را به دسته‌های مثبت، منفی یا خنثی تقسیم کنند. این نوع طبقه‌بندی در تحلیل احساسات، شناسایی موضوع و یا تفکیک انواع محتوا کاربرد دارد.

۲- پاسخ‌دهی به سؤالات (Question Answering - QA): مدل‌های نقاب‌دار می‌توانند به سؤالات مربوط به متون خاص پاسخ دهند. این کاربرد معمولاً به استخراج اطلاعات از یک متن یا مجموعه‌ای از متون مربوط می‌شود.

۳- شناسایی موجودیت‌های نام‌دار (Named Entity Recognition - NER): بدین صورت، مدل‌های زبانی نقاب‌دار، به شناسایی و طبقه‌بندی موجودیت‌های نام‌دار در متن مانند نام افراد، مکان‌ها، سازمان‌ها و... می‌پردازند. این مدل‌های زبانی می‌توانند با دقت بالا این موجودیت‌ها را شناسایی کرده و آن‌ها را در دسته‌های مختلف قرار دهند.

مدل‌های زبانی نقاب‌دار به‌عنوان یکی از پیشرفت‌های مهم در حوزه‌ی پردازش زبان طبیعی شناخته می‌شوند. این مدل‌ها با استفاده از تکنیک نقاب‌گذاری کلمات در جملات، قادر به پیش‌بینی و تولید متن‌های معنادار و با کیفیت بالا هستند. از جمله ویژگی‌های بارز این مدل‌ها



منابع:

- Jurafsky, D. (2000). Speech and language processing.
- Nozza, D., Bianchi, F., & Hovy, D. (2020). What the [mask]? making sense of language-specific BERT models. arXiv preprint arXiv:2003.02912.
- Yu, J., Kim, S. U., Choi, J., & Choi, J. D. (2024). What is your favorite gender, mlm? gender bias evaluation in multilingual masked language models. Information, 15(9), 549.
- Shah Jahan, M., Khan, H. U., Akbar, S., Umar Farooq, M., Gul, S., & Amjad, A. (2021). Bidirectional language modeling: a systematic literature review. Scientific Programming, 2021(1), 6641832.
- Chen, S., Peng, B., Chen, M., Wang, R., Xu, M., Zeng, X., ... & Lu, C. (2024). Causal Evaluation of Language Models. arXiv preprint arXiv:2405.00622.

می‌توان به توانایی آن‌ها در درک زمینه و ساختار زبان، تعمیم‌پذیری در وظایف مختلف زبانی و قابلیت‌های چندزبانی اشاره کرد. با وجود چالش‌هایی مانند نیاز به وجود داده‌های آموزشی زیاد و هزینه‌های محاسباتی بالا، مدل‌های زبان نقاب‌دار به‌عنوان ابزاری قدرتمند در توسعه برنامه‌های کاربردی متنوعی از جمله ترجمه‌ی ماشینی، تولید محتوا و تحلیل احساسات به شمار می‌روند.



نویسندگان:

محمدصادق پاکزاد - دانشجوی کارشناسی ارشد مهندسی کامپیوتر - گرایش هوش مصنوعی - ترم چهارم
آیدا میرزازاده - دانشجوی کارشناسی ارشد مهندسی کامپیوتر - گرایش هوش مصنوعی - ترم اول
پویان قنبری - دانشجوی کارشناسی ارشد مهندسی کامپیوتر - گرایش هوش مصنوعی - ترم دوم
آیرین پورعباس - دانشجوی کارشناسی ارشد مهندسی کامپیوتر - گرایش هوش مصنوعی - ترم دوم
مهتا مهرگانی - دانشجوی کارشناسی ارشد مهندسی کامپیوتر - گرایش هوش مصنوعی - ترم دوم



ماهیت زمانی زبان

زبان گفتاری و نوشتاری به صورت دنباله‌ای زمانی پردازش می‌شوند، به گونه‌ای که هر بخش از داده به ترتیب دریافت و پردازش می‌شود. این ماهیت زمانی در استعاره‌هایی مانند "جریان مکالمات" یا "فیدهای خبری" دیده می‌شود. الگوریتم‌های پردازش زبان مانند الگوریتم ویتربی، این خاصیت را منعکس می‌کنند و به صورت گام‌به‌گام پیش می‌روند. در مقابل، روش‌هایی مانند شبکه‌های عصبی پیش‌خور (Feedforward) یا دسته‌بندی متن فرض می‌کنند که به تمام ورودی به صورت هم‌زمان دسترسی دارند.

معرفی شبکه‌های عصبی بازگشتی (RNNs)

شبکه‌های عصبی بازگشتی (Recurrent neural networks) نوعی از شبکه‌های عصبی هستند که برای پردازش داده‌های دنباله‌ای طراحی شده‌اند. برخلاف شبکه‌های پیش‌خور که ورودی را مستقل از زمان پردازش می‌کنند، RNN ها می‌توانند ارتباط بین داده‌های متوالی را با استفاده از حافظه‌ی داخلی مدل‌سازی کنند.

RNN ها با اضافه کردن اتصالات بازگشتی به مدل، امکان پردازش دنباله‌ای و نگهداری اطلاعات در طول زمان را فراهم می‌کنند. ساختار آن‌ها به گونه‌ای است که حالت مخفی هر زمان، هم ورودی فعلی و هم حالت مخفی زمان قبل را در نظر می‌گیرد. به این ترتیب:

- حافظه‌ی کوتاه‌مدت RNN ها: قادرند اطلاعات اخیر را برای تصمیم‌گیری در هر لحظه ذخیره کنند.

- بدون محدودیت طول دنباله: برخلاف مدل‌های پنجره‌ای که تنها تعداد محدودی از کلمات قبل را می‌بینند، RNN می‌تواند اطلاعات کل دنباله را در نظر بگیرد.

زبان ماهیتی زمانی دارد و به صورت دنباله‌ای از رویدادها پردازش می‌شود. شبکه‌های عصبی بازگشتی (RNN) با استفاده از اتصالات بازگشتی و حافظه‌ی داخلی، توانایی مدل‌سازی داده‌های ترتیبی و پردازش آن‌ها را فراهم می‌کنند. معماری‌های پیشرفته‌تر (RNN) مانند LSTM، مشکلاتی نظیر گرادیان ناپدیدشونده و حساسیت به اطلاعات محلی را رفع می‌کنند و از مکانیزم‌هایی مانند گیت‌ها برای مدیریت حافظه استفاده می‌کنند. این مدل‌ها در کاربردهایی همچون مدل‌سازی زبان، برچسب‌گذاری دنباله‌ها، تولید متن و ترجمه‌ی ماشینی استفاده می‌شوند.

همچنین، معماری‌های کدگذار-رمزگشا (Encoder-Decoder) با مکانیزم توجه (Attention) محدودیت‌های مدل‌های سنتی را کاهش داده و توانایی تمرکز بر اطلاعات کلیدی ورودی را ارائه می‌دهند.

زبان یک پدیده‌ی ذاتاً زمانی است. زبان گفتاری به صورت دنباله‌ای از رویدادهای صوتی در طول زمان اتفاق می‌افتد و ما زبان گفتاری و نوشتاری را به صورت یک جریان ورودی پی‌درپی درک و تولید می‌کنیم. این ویژگی زمانی زبان در استعاره‌هایی که استفاده می‌کنیم، بازتاب یافته‌است؛ ما از جریان مکالمات، فیدهای خبری و جریان توییت‌ر صحبت می‌کنیم که همه نشان‌دهنده‌ی دنباله‌ای هستند که در طول زمان گسترش می‌یابد. شبکه‌ی عصبی بازگشتی و انواع آن مانند LSTM ها و RNN ها مکانیزمی دارند که مستقیماً با ماهیت پی‌درپی زبان سروکار دارد و به آن‌ها اجازه می‌دهد ویژگی زمانی زبان را بدون استفاده از پنجره‌های با اندازه‌ی ثابت مدیریت کنند. شبکه‌ی عصبی بازگشتی یک روش جدید برای نمایش زمینه‌های قبلی ارائه می‌دهد که از طریق اتصالات بازگشتی، تصمیم مدل را وابسته به اطلاعاتی از صدها کلمه‌ی قبلی می‌کند.

آموزش یک مدل زبانی RNN

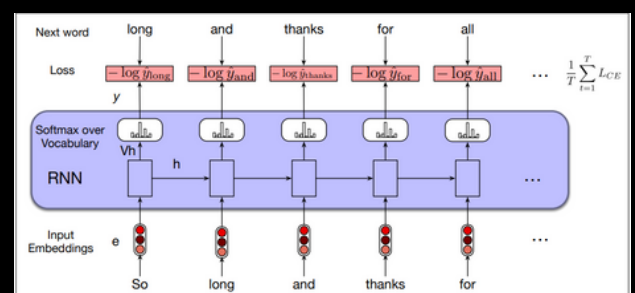
یک مجموعه‌ی متنی را به‌عنوان داده‌ی آموزشی در نظر می‌گیریم و در هر گام زمانی t از مدل می‌خواهیم تا کلمه‌ی بعدی را پیش‌بینی کند. به این مدل، مدل خودنظارتی می‌گوییم؛ زیرا مدل بدون نیاز به داده‌های برچسب‌دار خاص، فقط با خواندن متن و یادگیری از ترتیب کلمات، آموزش می‌بیند. ما مدل را به گونه‌ای آموزش می‌دهیم که خطا در پیش‌بینی کلمه‌ی بعدی در مجموعه‌ی آموزشی به حداقل برسد و برای این کار از تابع هزینه‌ی آنتروپی متقاطع (cross-entropy loss) استفاده می‌کنیم. این تابع میزان اختلاف میان خروجی «توزیع احتمال» پیش‌بینی‌شده با خروجی واقعی را اندازه‌گیری می‌کند.

$$L_{CE} = - \sum_w \tilde{y}_t[w] \log \hat{y}_t[w]$$

در مدل‌سازی زبان توزیع صحیح از دانستن کلمه‌ی بعدی به دست می‌آید. که برای این کار از بردار one-hot استفاده می‌کنیم به گونه‌ای که در آن مقدار مربوط به کلمه‌ی بعدی واقعی برابر با ۱ بوده و سایر مقادیر ۰ هستند. بنابراین تابع هزینه‌ی آنتروپی متقاطع برای مدل‌سازی زبان، بر اساس احتمالی که مدل به کلمه‌ی بعدی واقعی اختصاص می‌دهد، تعیین می‌شود که در زمان t برابر با منفی لگاریتم احتمال وجود کلمه‌ی بعدی در مجموعه‌ی آموزشی است.

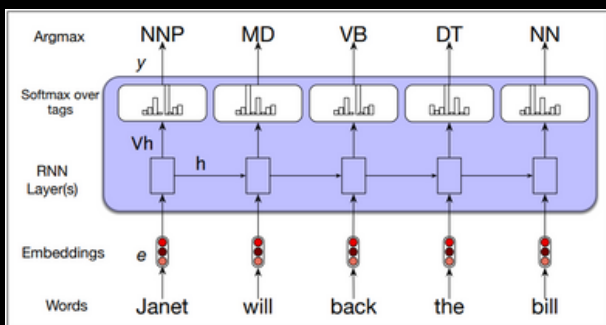
$$L_{CE}(\hat{y}_t, y_t) = - \log \hat{y}_t[w_{t+1}]$$

برای به حداقل رساندن خطای آنتروپی متقاطع در مجموعه‌ی آموزشی، وزن‌های شبکه را با استفاده از روش گرادینان نزولی تنظیم می‌کنیم.



برچسب‌گذاری دنباله‌ها

در برچسب‌گذاری دنباله‌ها، وظیفه‌ی مدل این است که به هر عنصر از دنباله، یک برچسب از مجموعه‌ی محدود و ثابتی از برچسب‌ها، اختصاص دهد. در روش RNN برای برچسب‌گذاری دنباله، ورودی‌های مدل، نمایش برداری از کلمات «تعبیه‌های کلمات» هستند و خروجی‌ها نیز، احتمال برچسب‌ها هستند که توسط یک لایه‌ی Softmax روی مجموعه‌ی داده‌شده از برچسب‌ها تولید می‌شوند به بیان ساده‌تر هر کلمه در یک دنباله به یک بردار عددی «تعبیه» تبدیل می‌شود و سپس شبکه‌ی عصبی با استفاده از این بردارها، احتمالی برای هر برچسب ممکن محاسبه می‌کند که در نهایت، برچسبی که بیشترین احتمال را داشته باشد به آن کلمه اختصاص داده می‌شود.



برچسب‌گذاری دنباله‌ها

در برچسب‌گذاری دنباله‌ها، وظیفه‌ی مدل این است که به هر عنصر از دنباله، یک برچسب از مجموعه‌ی محدود و ثابتی از برچسب‌ها، اختصاص دهد. در روش RNN برای برچسب‌گذاری دنباله، ورودی‌های مدل، نمایش برداری از کلمات «تعبیه‌های کلمات» هستند و خروجی‌ها نیز، احتمال برچسب‌ها هستند که توسط یک لایه‌ی Softmax روی مجموعه‌ی داده‌شده از برچسب‌ها تولید می‌شوند به بیان ساده‌تر هر کلمه در یک دنباله به یک بردار عددی «تعبیه» تبدیل می‌شود و سپس شبکه‌ی عصبی با استفاده از این بردارها، احتمالی برای هر برچسب ممکن محاسبه می‌کند که در نهایت، برچسبی که بیشترین احتمال را داشته باشد به آن کلمه اختصاص داده می‌شود.

ایجاد شود. با این روش، مدل می‌تواند اطلاعات گذشته و آینده را همزمان به دست آورد و از آن‌ها برای بهبود دقت در تحلیل توالی‌های زمانی یا زبانی استفاده کند.

مشکلات RNN ها

۱- گرادیان ناپدیدشونده (Vanishing Gradient): وقتی زنجیره‌ی دنباله طولانی باشد، سیگنال‌های گرادیان در حین آموزش کاهش یافته و اطلاعات دوردست از دست می‌روند.

۲- حساسیت به اطلاعات محلی: به دلیل معماری ساده، اطلاعات نزدیک بیشتر از اطلاعات دورتر در تصمیم‌گیری‌ها تأثیر می‌گذارند.

معرفی LSTM

برای پرداختن به این مسائل، معماری‌های شبکه پیچیده‌تری طراحی شده‌اند تا به‌طور صریح وظیفه‌ی حفظ بافت مربوطه را در طول زمان مدیریت کنند و شبکه را قادر می‌سازد تا اطلاعاتی را که دیگر مورد نیاز نیست فراموش کند و اطلاعات مورد نیاز برای تصمیم‌گیری‌های آینده را به خاطر بسپارد. متداول‌ترین مورد استفاده از چنین پسوندی برای RNN ها، شبکه‌ی حافظه‌ی کوتاه‌مدت-بلندمدت (LSTM یا Long short-term memory) است. LSTM ها مشکل مدیریت متن را به دو زیرمسئله تقسیم می‌کنند؛ حذف اطلاعاتی که دیگر مورد نیاز نیست و اضافه کردن اطلاعاتی که احتمالاً برای تصمیم‌گیری‌های بعدی مورد نیاز است.

کلید حل هر دو مشکل این است که یاد بگیرد چگونه این زمینه را مدیریت کنید تا اینکه یک استراتژی را در معماری سخت‌گذاری کنید. LSTM ها این کار را با افزودن یک لایه زمینه‌ی صریح به معماری «علاوه بر لایه‌ی پنهان معمولی تکرارشونده» و از طریق استفاده از واحدهای عصبی تخصصی که از گیت‌ها برای کنترل جریان اطلاعات به داخل و خارج از واحدها استفاده می‌کنند، انجام می‌دهند. این گیت‌ها از طریق استفاده از وزنه‌های اضافی که به صورت متوالی بر روی ورودی و لایه‌ی پنهان قبلی و لایه‌های زمینه‌ی قبلی عمل

همان‌طور که در شکل بالا مشاهده می‌کنید، ورودی‌ها در هر گام زمانی، تعبیه‌های از پیش آموزش‌دیده‌شده‌ی کلمات متناظر با توکن‌های ورودی هستند.

بلوک RNN در تصویر، نمایش ساده‌شده‌ای از یک شبکه‌ی بازگشتی را نشان می‌دهد که شامل یک لایه‌ی ورودی، یک لایه‌ی پنهان، یک لایه‌ی خروجی در هر زمان و ماتریس‌های مشترک V و U و ماتریس وزن W است که شبکه را تشکیل می‌دهند.

خروجی‌های شبکه در هر زمان، توزیع احتمال برچسب‌های بخش گفتار را نشان می‌دهند که توسط لایه‌ی Softmax محاسبه می‌شوند. جهت تولید دنباله‌ای از برچسب‌ها برای یک ورودی مشخص، روش استدلال مستقیم (Forward inference) را بر روی دنباله‌ی ورودی اجرا کرده و محتمل‌ترین برچسب را از Softmax در هر گام زمانی انتخاب می‌کنیم.

از آنجایی که از لایه‌ی Softmax برای تولید توزیع احتمال بر روی مجموعه‌ی برچسب‌های خروجی استفاده می‌کنیم، مجدداً از تابع هزینه‌ی آنتروپی متقاطع در طول آموزش استفاده خواهیم کرد.

معماری‌های شبکه‌ی عصبی بازگشتی پشته‌ای و دوطرفه به معماری‌های خاصی از شبکه‌های عصبی بازگشتی اشاره دارد که برای مدل‌سازی بهتر دنباله‌ها و اطلاعات ترتیبی در داده‌ها طراحی شده‌اند.

۱- شبکه‌های عصبی بازگشتی پشته‌ای (Stacked RNNs): این معماری به این معناست که چندین لایه RNN روی هم قرار می‌گیرند و هر لایه‌ی خروجی لایه‌ی قبلی را به‌عنوان ورودی خود دریافت می‌کند. هدف از این پیکربندی، یادگیری ویژگی‌های پیچیده‌تر از داده‌های ترتیبی است.

۲- شبکه‌های عصبی بازگشتی دوطرفه (Bidirectional RNNs): در این نوع معماری دو RNN مجزا روی یک دنباله پردازش می‌شوند، یکی در جهت رو به جلو و دیگری در جهت رو به عقب. خروجی‌های این دو RNN در هر زمان ترکیب می‌شوند تا یک نمایش کامل‌تر از دنباله

در مرحله‌ی بعد، ماسکی را برای افزودن دروازه ایجاد می‌کنیم تا اطلاعاتی را برای افزودن به زمینه‌ی فعلی انتخاب کنیم.

$$i_t = \sigma(U_i h_{t-1} + W_i x_t)$$

$$\tilde{i}_t = g_i \odot i_t$$

در مرحله‌ی بعد، این را به بردار زمینه‌ی اصلاح‌شده اضافه می‌کنیم تا بردار زمینه‌ی جدید خود را به دست آوریم.

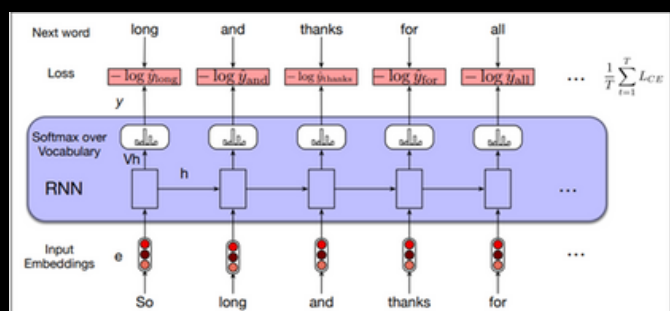
$$c_t = \tilde{i}_t + k_t$$

دروازه نهایی «خروجی» که ما استفاده خواهیم کرد، دروازه‌ی خروجی است که برای تصمیم‌گیری درباره‌ی اینکه چه اطلاعاتی برای وضعیت پنهان فعلی مورد نیاز است «برخلاف اطلاعاتی که باید برای تصمیم‌گیری‌های آینده حفظ شود» استفاده می‌شود.

$$o_t = \sigma(U_o h_{t-1} + W_o x_t)$$

$$h_t = o_t \odot \tanh(c_t)$$

دروازه نهایی «خروجی» که ما استفاده خواهیم کرد، دروازه‌ی خروجی است که برای تصمیم‌گیری درباره‌ی اینکه چه اطلاعاتی برای وضعیت پنهان فعلی مورد نیاز است «برخلاف اطلاعاتی که باید برای تصمیم‌گیری‌های آینده حفظ شود» استفاده می‌شود.



می‌کنند، اجرا می‌شوند. دروازه‌ها در یک LSTM یک الگوی طراحی مشترک دارند. هر یک از یک لایه‌ی پیش‌خور تشکیل شده‌است، به‌دنبال آن یک تابع فعال‌سازی سیگموئید، و به‌دنبال آن یک ضرب نقطه‌ای با لایه‌ای که درجه شده‌است. انتخاب سیگموئید به‌عنوان تابع فعال‌سازی از تمایل آن به فشار دادن خروجی‌های خود به ۰ یا ۱ ناشی می‌شود. ترکیب آن با ضرب نقطه‌ای تأثیری شبیه به یک ماسک باینری دارد. مقادیر موجود در لایه‌ی در حال دروازه که با مقادیر نزدیک به ۱ در ماسک تراز می‌شوند تقریباً بدون تغییر منتقل می‌شوند. مقادیر مربوط به مقادیر پایین‌تر اساساً پاک می‌شوند.

اولین دروازه‌ای که در نظر خواهیم گرفت دروازه‌ی فراموشی است. هدف این گیت حذف اطلاعاتی از زمینه است که دیگر مورد نیاز نیست. دروازه‌ی فراموشی مجموع وزنی لایه‌ی پنهان حالت قبلی و ورودی فعلی را محاسبه می‌کند و آن را از یک سیگموئید عبور می‌دهد. سپس این ماسک بر حسب عنصر در بردار زمینه ضرب می‌شود تا اطلاعاتی که دیگر مورد نیاز نیست حذف شوند. ضرب عنصری دو بردار «که توسط عملگر نشان داده می‌شود و گاهی اوقات حاصل ضرب هادامارد نیز نامیده می‌شود» همان ابعاد دو بردار ورودی است که در آن هر عنصر حاصل ضرب عنصر i در دو بردار ورودی است:

$$f_t = \sigma(U_f h_{t-1} + W_f x_t)$$

$$k_t = c_{t-1} \odot f_t$$

وظیفه‌ی بعدی محاسبه‌ی اطلاعات واقعی است که باید از حالت پنهان قبلی و ورودی‌های فعلی استخراج کنیم؛ همان محاسبات اولیه‌ای که برای همه‌ی شبکه‌های تکراری خود استفاده می‌کردیم.

$$g_t = \tanh(U_g h_{t-1} + W_g x_t)$$

کاربردهای RNN و LSTM

۱- مدل سازی زبان:

الف) پیش بینی کلمه ی بعدی در دنباله ی متنی.
ب) برای مثال، در جمله ی «من یک سیب می خواهم»، مدل احتمال کلمه ی بعدی را با توجه به کلمات قبل محاسبه می کند.

۲- برچسب گذاری دنباله ها:

الف) وظایفی مانند برچسب گذاری نقش های دستوری (POS tagging) یا تشخیص موجودیت های نامدار.
ب) مدل در هر گام به هر کلمه یک برچسب خاص اختصاص می دهد.

۳- دسته بندی دنباله ها:

الف) وظایفی مانند تحلیل احساسات یا طبقه بندی موضوعی.
ب) مدل تمام دنباله را بررسی کرده و در پایان یک کلاس به آن اختصاص می دهد.

۴- تولید متن:

الف) تولید خودکار متن برای وظایفی مانند ترجمه ماشینی، خلاصه سازی، یا ایجاد داستان.
ب) مدل به صورت گام به گام کلمات جدید را تولید می کند.

۵- مدل های کدگذار-رمزگشا (Encoder-Decoder):

الف) تبدیل دنباله های با طول متفاوت مانند ترجمه ی زبان ها.
ب) کدگذار ورودی را به یک بردار فشرده تبدیل می کند و رمزگشا این بردار را به خروجی تبدیل می کند.

مدل کدگذار-کدگشا با RNN ها

هدف کلی از کدگذار این است که یک نمایش متنی (یا زمینه ای) از ورودی تولید کند. این نمایش در وضعیت پنهانی نهایی کدگذار نمایان می شود hen. این نمایش،

که به آن C برای زمینه (Context) نیز گفته می شود، سپس به کدگشا منتقل می شود. ساده ترین نسخه از شبکه کدگشا، این وضعیت را گرفته و از آن تنها برای مقداردهی اولیه به اولین وضعیت پنهانی کدگشا استفاده می کند؛ اولین سلول RNN کدگشا از C به عنوان وضعیت پنهانی اولیه h0d خود استفاده می کند. سپس کدگشا به طور خودبازگشتی یک دنباله از خروجی ها را تولید می کند؛ هر بار یک عنصر، تا زمانی که نشانگر پایان دنباله تولید شود. هر وضعیت پنهانی بر اساس وضعیت پنهانی قبلی و خروجی تولید شده در وضعیت قبلی مشروط می شود. به طور خلاصه:

- برای وظایفی که ورودی و خروجی طول متفاوت دارند مانند ترجمه ی زبان ها استفاده می شود.

- کدگذار: اطلاعات ورودی را به یک بردار زمینه فشرده می کند.

- رمزگشا: بردار زمینه را به خروجی تبدیل می کند.

- این معماری پایه ی مدل های پیشرفته ای چون ترانسفورمرها می باشد.

- آموزش مدل کدگذار-کدگشا در بخش.

معماری های کدگذار-کدگشا به صورت انتها به انتها (end to end) آموزش داده می شوند. هر مثال آموزشی یک جفت رشته متصل به هم است، یکی به عنوان منبع و دیگری به عنوان هدف. با افزودن یک توکن جداکننده، این جفت های منبع-هدف اکنون می توانند به عنوان داده های آموزشی مورد استفاده قرار گیرند.

برای ترجمه ی ماشینی، داده های آموزشی معمولاً شامل مجموعه هایی از جملات و ترجمه های آن ها است. این داده ها می توانند از مجموعه های استاندارد جفت جملات همراستا استخراج شوند، وقتی که مجموعه ی داده ی آموزشی را داشته باشیم، فرایند آموزش همان طور که در هر مدل زبانی مبتنی بر RNN انجام می شود، پیش می رود. شبکه متن ورودی «منبع» را دریافت می کند و سپس با شروع از توکن جداکننده به طور خودبازگشتی آموزش داده می شود تا کلمه ی بعدی را پیش بینی کند.



منابع:

- 1- Jurafsky, D., & Martin, J. H. (2024). Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition.
- 2- Jurafsky, D. (2000). Speech and language processing.

مکانیزم توجه

- با ارائه‌ی بردارهای وزن‌دهی‌شده، مدل می‌تواند بر بخش‌های مهم‌تر ورودی تمرکز کند.
- مزیت اصلی: حل مشکل محدودیت بردار زمینه در مدل‌های کدگذار-رمزگشا.

شبکه‌های عصبی بازگشتی و معماری‌های مشتق از آن، ابزارهای کارآمدی برای مدل‌سازی داده‌های ترتیبی و زبانی هستند. معرفی مکانیزم‌هایی مانند RNN، LSTM و Attention موجب بهبود قابلیت پردازش اطلاعات طولانی‌مدت و افزایش دقت در وظایف پردازش زبان طبیعی شده‌است. این پیشرفت‌ها زمینه‌ساز توسعه‌ی مدل‌های مدرن مانند ترانسفورمرها و بهبود عملکرد در کاربردهای پیچیده‌ای مانند ترجمه‌ی ماشینی و تولید متن شده‌اند.



نویسندگان:

- حوری طالش امیری - دانشجوی کارشناسی ارشد مهندسی کامپیوتر - گرایش هوش مصنوعی - ترم سوم
- زهرا قنبریور - دانشجوی کارشناسی ارشد مهندسی کامپیوتر - گرایش هوش مصنوعی - ترم اول
- فاطمه گل محمدی - دانشجوی کارشناسی ارشد مهندسی کامپیوتر - گرایش هوش مصنوعی - ترم دوم
- محمد امید پور دانشجوی کارشناسی ارشد مهندسی کامپیوتر - گرایش هوش مصنوعی - ترم دوم

گیلانو

  SCC_LIAU  scc.liau@gmail.com

گیلانو نشریه‌ای دانشجویی در زمینه علمی تخصصی با صاحب امتیازی انجمن علمی مهندسی کامپیوتر دانشگاه آزاد اسلامی واحد لاهیجان است که با ترتیب انتشار ماهانه منتشر می‌گردد.

گیلانو از اردیبهشت ۱۴۰۰ شروع به فعالیت کرد و در آن به موضوعات مرتبط با تمام گرایش‌های مهندسی کامپیوتر نظیر هوش مصنوعی، رباتیک، نرم‌افزار، سخت‌افزار، شبکه و موضوعات بین رشته‌ای پرداخته می‌شود.

هیات تحریریه گیلانو شامل دانشجویانی از رشته‌هایی مختلفی چون مهندسی کامپیوتر، مهندسی برق، مهندسی پزشکی میکروبیولوژی، روانشناسی، پرستاری و... است.