جزوه آموزشی جامع: مبانی هوش مصنوعی و فنون تعامل هوشمند "جلسه اول"

این جزوه آموزشی با هدف اصلاح مدل ذهنی پژوهشگران در مواجهه با ابزارهای نوین و آموزش استانداردهای تعامل حرفه‌ای با مدل‌های زبانی بزرگ، بر اساس مباحث مطرح شده در کلاس تدوین شده است.

۱. ماهیت هوش مصنوعی: رفتار عقلایی در مقابل محاسبات آماری

نخستین گام در شناخت هوش مصنوعی، درک این حقیقت است که آنچه ما «هوش» می‌نامیم، در واقع یک «شبیه‌سازی بسیار پیشرفته از رفتار عقلایی» است، نه وجود یک عقل ادراک‌گر.

۱. **ریشه رفتار :**رفتار انسان منشأ تجربی و شهودی دارد، اما پاسخ‌های هوش مصنوعی صرفاً حاصل محاسبات ریاضی بر روی توده‌های عظیم داده است.
۲. **چالش اعتماد :**در فرایندهای حساس پژوهشی "مانند نقد، استدلال و تحلیل"، نمی‌توان به هوش مصنوعی به عنوان یک «موجود صاحب عقل» اتکا کرد؛ بلکه باید با «تزریق تعقل انسانی» به ماشین، خروجی آن را مدیریت کرد.
۳. **ماشین‌های قدیمی در برابر مدل‌های مولد :**تفاوت اصلی در «مولد بودن» است. ماشین‌های قدیمی طبق برنامه از پیش تعیین شده عمل می‌کردند، اما مدل‌های امروزی بر اساس احتمال، پاسخ را در لحظه «خلق» می‌کنند.

۲. استعاره اسب عربی: مدل ذهنی صحیح برای استفاده از AI

برای تعامل با هوش مصنوعی مولد، باید مدل ذهنی خود را از «ماشین‌حساب» به «موجود زنده قوی اما بی‌اختیار» تغییر دهید. هوش مصنوعی مانند یک **«اسب اصیل و نجیب عربی»** است:

- بسیار قوی، سریع و خیره‌کننده است.
- بدون سوارکار ماهر، ممکن است رم کند یا سوار را به زمین بکوبد.
- **سوارکاری "مهارت پژوهشگر" :**شناخت محدودیت‌های این ابزار، همان هنر سوارکاری است. شما باید بدانید کجا افسار را بکشید و کجا اجازه تاختن بدهید.

«هوش مصنوعی یک ابزار است و مسئولیت نهایی پژوهش بر عهده پژوهشگر است».

۳. توهم زنده پنداری و خطای تعامل عاطفی

یکی از بزرگترین موانع در پژوهش دیجیتال، «توهم زنده‌پنداری "Sentience Illusion" «است.

- **عبرت تاریخی :**زمانی که رادیو وارد جوامع سنتی شد، برخی "مانند سلفی‌ها و اهل‌سنت در آن مقطع" گمان می‌کردند صدای اجنه است، چون صدا بود اما صاحب صدا دیده نمی‌شد. امروز نیز برخی یا با «شیطان‌انگاری» هوش مصنوعی را رد می‌کنند، یا با «پذیرش چشم‌بسته»، به آن شخصیت انسانی می‌بخشند.
- **تعامل مکانیکی :**هوش مصنوعی نه ادب می‌فهمد و نه احساس. برای گرفتن بهترین نتیجه، باید از «ارتباط عاطفی» پرهیز کرده و به «ارتباط مکانیکی» روی بیاورید.

چک‌باکس دستورالعمل تعامل حرفه‌ای:

- [ ]استفاده از جملات صریح و امری "مانند: «استخراج کن»، «فیش‌برداری کن»، «نقد کن."«
- [ ]اجتناب از مقدمه‌چینی‌های عاطفی و تعارفات انسانی.
- [ ]تمرکز بر «تزریق تعقل» از طریق دستورات دقیق ساختاری.

٤ .کالبدشناسی مدل‌های بزرگ زبانی "LLM" و مکانیسم پیش‌بینی

مدل‌های بزرگ زبانی مانندGPT ، در واقع مخازن عظیمی هستند که حدود **۳۷۰ میلیارد پارامتر** "یا بیشتر" را در خود جای داده‌اند. فرایند پاسخ‌دهی آن‌ها نه «تفکر»، بلکه «چیدن کلمات بر مبنای محاسبات ریاضی و آماری» است.

- **مثال اول :**وقتی عبارت «الحمدالله رب...» را می‌شنوید، ذهن شما "و هوش مصنوعی" بر اساس تکرار بی‌شمار، کلمه «العالمین» را پیش‌بینی می‌کند.
- **مثال دوم :**در عبارت «پیش از نماز احمد...»، مدل پیش‌بینی می‌کند که احتمال کلمه «وضو گرفت» ۸۰٪ و احتمال «اذان گفت» یا «نیت کرد» بسیار کمتر است.

⚠ بسیار مهم :هوش مصنوعی همواره «محتمل‌ترین» کلمه را بر اساس الگوهای آماری کنار هم می‌چیند، نه لزوماً «صحیح‌ترین» یا «واقعی‌ترین» مطلب را. او برای جلب رضایت شما طراحی شده است، نه برای صیانت از حقیقت.

٥ .پدیده توهم "Hallucination" و جعل داده درAI

هوش مصنوعی درکی از «کذب» و «دروغ» ندارد. او فقط «قالب‌ها» را یاد گرفته است.

- **جعل شیک و متقاعدکننده :**مدل می‌داند که یک حدیث معتبر باید با «قال الصادق "ع"» شروع شود و متن عربی با اعراب‌گذاری داشته باشد. لذا ممکن است حدیثی بسازد که در هیچ منبعی وجود ندارد اما ساختاری کاملاً حرفه‌ای دارد.
- **جعل ترکیبی :**در مباحث تخصصی، مدل ممکن است داده‌های مختلف را با هم ترکیب کند "مانند نسبت دادن نظریه اعتبارات علامه طباطبایی به شخصی دیگر" تا فقط به شما پاسخ داده باشد.

توصیه راهبردی "فتوای آموزشی" :هرگز و تحت هیچ شرایطی از چت‌بات‌های عمومی برای استخراج و استناد به نصوص مقدس "آیات قرآن و احادیث" استفاده نکنید. ضریب جعل و توهم در این موارد به دلیل حساسیت کلمات بالاست و ابزار فعلاً قابل اتکا نیست.

۶. تکنیک فیوشات "Few-Shot" و مدیریت پرامپت

بهترین راه برای مهار توهم هوش مصنوعی، آموزش مدل در حین پرامپت‌نویسی از طریق ارائه مثال است. این کار باعث می‌شود مدل «الگوی ذهنی» شما را بفهمد.

- **مثال‌های ایجابی :**نمونه‌هایی از آنچه می‌خواهید.
- **مثال‌های سلبی :**نمونه‌هایی از آنچه نمی‌خواهید یا خط قرمز شماست.

نمونه کاربردی پرامپت "قابل کپی و استفاده":

»من می‌خواهم نظر علما را در مورد "حکم اقامت ۱۰ روزه مسافر" استخراج کنی. به این الگو توجه کن: ۱. آیت‌الله سیستانی: [متن نظر]، منبع: [نام کتاب]. "مثال ایجابی ۱" ۲. آیت‌الله مکارم شیرازی: [متن نظر]، منبع: [نام کتاب]. "مثال ایجابی ۲" ۳. آیت‌الله نوری همدانی: در این مورد هیچ نظری از ایشان در منابع یافت نشد. "مثال سلبی برای آموزش محدودیت" حالا با همین الگو و رعایت دقیق منابع، نظر [شخص مورد نظر] را استخراج کن«.

۷. روش‌های پیشرفته کنترل خطا RAG" و کانتکست"

برای اینکه هوش مصنوعی «جعل ترکیبی نکند»، باید حافظه آماری او را محدود کرده و او را به یک منبع مشخص متصل کنید.

- **تکنولوژی RAG "بازیابی تقویت‌شده" :**در این روش، شما کتاب یا متن مشخصی را به مدل می‌دهید و می‌گویید: «فقط و فقط از این متن پاسخ بده.«
- **استعاره سعدی و شاهنامه :**اگر از حافظه مدل بپرسید فلان شعر متعلق به کیست، ممکن است به دلیل شباهت سبک، شعر سعدی "گلستان" را به شاهنامه نسبت دهد. اما باRAG ، شما کتاب گلستان را باز کرده و جلوی روی او می‌گذارید تا از روی آن بخواند.
- **تفاوت ابزاری :**برای جستجوی مستقیم در وب و منابع جدید، ابزارهای تخصصی مانند **Perplexity** بسیار دقیق‌تر از چت‌بات‌های معمولی "مثل نسخه رایگان ChatGPT" عمل می‌کنند.

گام‌های طلایی برای مهار توهم :۱. دستورات را به صورت امری و مرحله‌بندی شده بدهید. ۲. از تکنیک **Few-Shot** "حداقل ۳ مثال" استفاده کنید. ۳. منبع پاسخگویی را با بارگذاری فایل "Context" محدود کنید. ۴. صراحتاً بنویسید: «اگر پاسخ در متن نبود، بگو نمی‌دانم و از خودت چیزی نساز.«

۸. جمع‌بندی: شما «سرمربی» هستید

در نهایت، هوش مصنوعی یک بازیکن است و شما **سرمربی** "Head Coach" هستید. در دنیای حرفه‌ای، اگر تیم شکست بخورد، کسی بازیکن را بازخواست نمی‌کند؛ این سرمربی است که در کنفرانس خبری باید پاسخگو باشد و مسئولیت را گردن بگیرد.

هوش مصنوعی می‌تواند در تلخیص، فیش‌برداری و روان‌سازی متن به شما کمک کند، اما اعتبار علمی پژوهش و صحت‌سنجی نهایی ارجاعات، وظیفه تخطی‌ناپذیر پژوهشگر است. هوش مصنوعی بدون نظارت شما، تنها یک «ماشین تولید احتمال» است.