

فایل توضیحات پروژه پایانی درس مدل‌های زبانی بزرگ

اعضای گروه:

بردیا خلفی - ۸۱۰۱۹۹۴۱۴ | مجید فریدفر - ۸۱۰۱۹۹۵۶۹ | علی آذر سینا - ۸۱۰۱۰۲۰۶۷

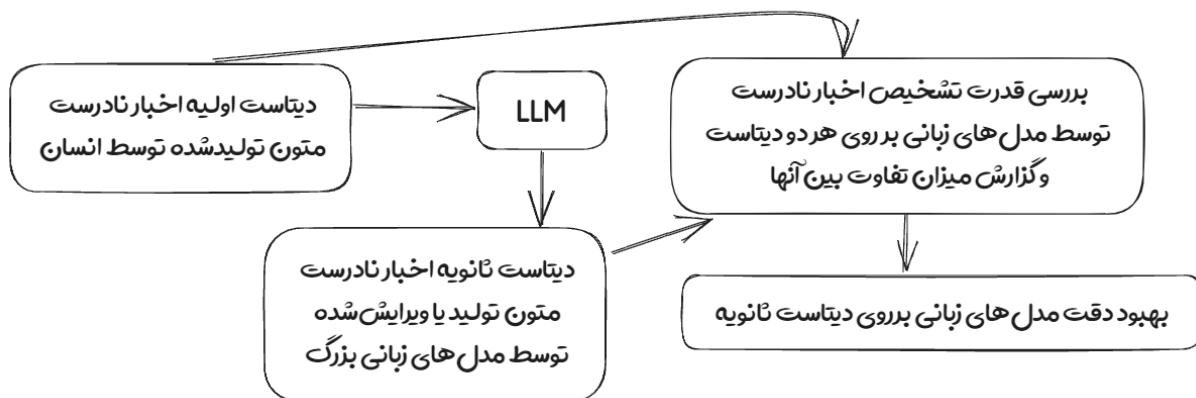
عنوان پروژه:

شناسایی و تشخیص اخبار نادرست (fake-news) تولید شده توسط مدل‌های زبانی بزرگ

توضیحات پروژه:

مسئله تشخیص اخبار نادرست و خودکارسازی این فرآیند، همواره یکی از مسائل چالشی برای بشریت بوده است و کارهای زیادی در این زمینه انجام شده است. در طی سال‌های اخیر و با ظهور مدل‌های زبانی بزرگ، چالش جدیدی ایجاد شده است؛ کاربران استفاده‌کننده از این مدل‌ها قادر هستند متن اخبار نادرست خود را با کمک این مدل‌ها ویرایش کنند تا متنی قابل قبول و غیرقابل تشخیص ایجاد کنند. گزارش‌ها نشان می‌دهد که تشخیص خودکار نادرستی این نوع اخبار بسیار سخت‌تر از اخبار کاملاً ساخته‌شده به دست انسان است.

در این پروژه قصد داریم تا در ابتدا به وسیله یک دیتاست اولیه از اخبار نادرست و با کمک مدل‌های زبانی بزرگ، دیتاست ثانویه‌ای از اخبار نادرست را ایجاد کنیم. سپس مجدداً با کمک مدل‌های زبانی بزرگ، ارزیابی از میزان دقت این مدل‌ها در تشخیص اخبار نادرست هر دو دیتاست داشته باشیم و میزان تفاوت را بررسی کنیم. سپس در ادامه با کمک تکنیک‌های یادگرفته شده در درس، به بهبود کیفیت و دقت این مدل‌ها در تشخیص اخبار نادرست بپردازیم.



شکل ۱ - نمودار توضیحات کلی پروژه