⚡️🚀המאמר היומי של מייק 10.09.24: ⚡️🚀

Can LLMs Generate Novel Research Ideas? A Large-Scale Human Study with 100+ NLP Researchers

האם מודלי שפה יכולים לייצר רעיונות מחקר חדשניים? 🤔 מחקר חדש מעורר גלים. ראינו לאחרונה התלהבות רבה סביב השימוש ב-LLMs לגילויים מדעיים. אך האם הם באמת מסוגלים להגיע לרעיונות חדשניים ברמת ראויה לחוקר במוסד אקדמי או בתעשיה?

מחברי המאמר תכננו ניסוי כדי לבדוק את הסיפור הזה. הם שכרו מעל 100 מומחי עיבוד שפה טבעית לכתוב רעיונות מחקר ולבחון רעיונות שנוצרו על ידי בני אדם ו-LLMs (בעיוור כלומר הבודקים לא ידעו מה מקורה של הרעיון שהם בודקים).

מתברר כי הרעיונות של ה-LLM נשפטו (באופן לא מפתיע קלוד נבחר למשימה זו) כחדשניים יותר מרעיונות מומחים אנושיים (עם מובהקות סטטיסטית), אך דורגו נמוך יותר בהיתכנות.

המחברים מציינים את מהחוזקות הבאות של רעיונות ה-LLM:

- הצעת מכילה שילובים ייחודיים של טכניקות מדומיינים שונים

- חקירת תחומים שלא נחקרו מספיק

- יצירת ניסויי מחשבה יצירתיים ומקוריים

עם זאת, היו להם גם כמה נקודות בעיותיות:

- חוסר פירוט מספק בנוגע ליישום

- שימוש לא נכון במאגרי נתונים

- החמצת בייסליינים (לא מפתיע כלל)

- הנחות לא מציאותיות

לעומת זאת, רעיונות אנושיים נטו להיות מעוגנים יותר במחקר קיים ובשיקולים מעשיים, אך לעתים קרובות היו פחות חדשניים, ובנו באופן הדרגתי על אינטואיציות ותוצאות ידועות.

המחברים מציינים שהחוקרים מכירים בקושי לשפוט חדשנות, אפילו עבור מומחים. כצעד הבא, הם הציעו לתת לחוקרים לממש את הרעיונות הללו, כדי לראות אם דירוגי החדשנות וההיתכנות מתורגמים להבדלים משמעותיים במציאות.

https://arxiv.org/abs/2409.04109