⚡️🚀המאמר היומי של מייק -23.11.24: ⚡️🚀

Table Meets LLM: Can Large Language Models Understand Structured Table Data? A Benchmark and Empirical Study

היום אני סוקר מאמר בנושא שמזמן לא נגעת בו(בסקירות) והוא דאטה טבלאי. המאמר בוחן שאלה מרתקת - האם מודלי שפה גדולים (LLMs) כמו GPT באמת מבינים מידע מובנה בטבלאות?

## קצת: רקע

בשנים האחרונות, LLMs הפכו לכלי חשוב בעיבוד שפה טבעית. אבל בעוד שהם מצוינים (סוג של) בהבנת שפה טבעית (בצורה של טקסט), יכולתם להבין מידע בצורה של טבלאות עדיין לא נחקרה לעומק וזה בדיוק מה שהחוקרים מנסים לעשות במאמר המסוקר

## מה החוקרים עשו?

החוקרים פיתחו מדד חדש שנקרא (SUC (Structural Understanding Capabilities שבוחן את היכולות של מודלים להבין מבנה של טבלאות. המדד כולל שבע משימות שונות:

זיהוי גבולות טבלה

איתור תאים ספציפיים

חיפוש הפוך (מיקום לערך)

אחזור עמודות

אחזור שורות

זיהוי גודל טבלה

זיהוי תאים ממוזגים

הם בדקו את GPT-3.5 ו-GPT-4 במשימות אלו תוך שימוש בפורמטים שונים של קלט (HTML, JSON, CSV ועוד).

## מה הם גילו?

התוצאות מפתיעות! הנה הנקודות העיקריות:

HTML מתגלה כפורמט ״הנוח״ ביותר להצגת טבלאות ל-LLMs

המודלים הראו יכולות טובות במשימות יחסיות מורכבות ( זיהוי גבולות טבלה, זיהוי תאים ממוזגים) אך נכשלו במשימות פשוטות ( זיהוי גודל טבלה, אחזור שורה פשוט, חיפוש תא בודד)

הביצועים השתפרו משמעותית עם דוגמה אחת (one-shot) לעומת אפס דוגמאות

## החידוש המרכזי: Self-augmented Prompting

החוקרים פיתחו שיטה חדשה שנקראת "self-augmented prompting" שמשפרת את ביצועי המודלים. השיטה מבקשת מהמודל תחילה לזהות מידע קריטי בטבלה (כמו טווחי ערכים) ואז משתמשת במידע הזה כדי לשפר את התשובה הסופית. זה מאפשר שיפור די רציני במספר בנצ'מארקים)

## סיכום:

אני חייב להגיד שהמאמר הזה מרתק. הוא מראה שלמרות ההתקדמות העצומה ב-LLMs, יש עדיין פערים משמעותיים ביכולת שלהם להבין מידע מובנה. זה מזכיר לנו שלמרות שהמודלים האלה מרשימים, הם עדיין רחוקים מהבנה אנושית אמיתית של מבנים ויחסים בין דאטה.

החוקרים עשו עבודה לא רעה בפיתוח מדדים ושיטות שיעזרו לקהילה להמשיך לשפר את היכולות האלה. השיטה החדשה שלהם ל-prompting היא פשוטה אבל אפקטיבית, וזה בדיוק מה שאנחנו צריכים - פתרונות פרקטיים שאפשר ליישם מיד.

## מילה אחרונה

אם אתם עובדים עם טבלאות ו-LLMs, המאמר הזה הוא חובה. הוא מספק תובנות מעשיות וכלים שימושיים. הקוד והדאטה זמינים ב-GitHub, אז אתם יכולים להתחיל לשחק עם זה ישר.

מעניין במיוחד יהיה לראות איך הממצאים האלה ישפיעו על הדור הבא של מודלי שפה. האם נראה מודלים שמתוכננים במיוחד להבנת מידע מובנה?

https://arxiv.org/abs/2305.13062