המאמר היומי של מייק: 17.05.25  
ZEROSEARCH: Incentivize the Search Capability of LLMs without Searching

במאמר ZEROSEARCH מוצגת שיטה חדשה לאימון יכולות חיפוש של מודלי שפה גדולים בעזרת שיטת מלמידה עם חיזוקים (RL), מבלי להשתמש כלל במנועי חיפוש אמיתיים. במקום לגשת ל-Google או ל-API חיצוני, הם מאמנים מודל שפה קטן יותר לשמש כסימולציית מנוע חיפוש, שמספק מסמכים רלוונטיים או רועשים לפי צורך.

המודל המדמה עובר Fine-tuning על זוגות של שאילתות-תשובות מתוך אינטראקציות אמיתיות עם מנוע חיפוש, מסמכים שהובילו לתשובות נכונות מסומנים כחיוביים, ואחרים כשליליים(זה די מקורי האמת כי עושים זאת בד"כ עם RL). לצורך כך, הם שומרים גם את השאלה המקורית והתשובה הנכונה בתוך הפרומפט, כדי לאפשר למודל ללמוד הקשרים סמנטיים עמוקים יותר ולשלוט באיכות התשובה. לאחר מכן, במהלך האימון ב-RL, המודל המרכזי מייצר שאילתות חיפוש, מקבל את המסמכים מהמודל המדמה, ומבצע reasoning כדי להפיק תשובה.

כדי לחזק את יכולות ההסקה, הם מוסיפים מנגנון לימוד מדורג: איכות המסמכים יורדת לאורך זמן האימון באופן מבוקר, מה שמכריח את המודל להתמודד עם מידע חלקי או שגוי. האימון נעשה בעזרת אלגוריתמים כמו PPO ו-GRPO, שזה דווקא די סטנדרטי. התוצאה: מערכת שמגיעה ואף עוברת ביצועים של מודלים המשתמשים בגוגל, ללא עלות API ובשליטה מלאה באיכות המידע.

מאמר נחמד על איך לחפש ללא חיפוש באמצעות llms.

https://arxiv.org/abs/2505.04588