המאמר היומי של מייק: 30.05.25  
Learn Beyond the Answer: Training Language Models with Reflection for Mathematical Reasoning

בדיוק לפני שנה (30.05.24) התחלתי את הסקירה היומית ופתחתי ערוץ טלגרם Science and AI with Mike בשבילם. מאז כתבתי 253 סקירות שזה עושה סקירה ל 1.44 יום למרות שלאחרונה קצת האטתי את הקצב.

אז לרגל תאריך התחלתי לסקור מאמר בן קצת פחות משנה בנושא חשוב…

המאמר מציג חידוש מושגי ומתודולוגי חשוב (נכון ללפני 10 חודשים לאימון מודלים בשיקול דעת מתמטי. בניגוד לרוב המאמרים שיצאו לפניו, שמתמקדים בהעשרת הדאטה באמצעות שאלות חדשות או תשובות שונות (augmentations על גבי מימד הדוגמאות), החידוש המרכזי כאן הוא בהצגת גישה חדשה לגמרי: הרחבה רפלקטיבית (Reflective Augmentation).

במקום להוסיף עוד דוגמאות, המחברים מציעים להעמיק כל דוגמה קיימת ע"י הוספת מקטע רפלקטיבי אחרי הפתרון הסטנדרטי, הכולל שני רכיבים עיקריים:

מהלך אלטרנטיבי – פתרון נוסף שמציע נקודת מבט שונה על אותה בעיה.

הרחבה (Follow-up) – ניסוח בעיה כללית או אנלוגית שמעמיקה את ההבנה.

החדשנות אינה רק טכנית אלא קוגניטיבית: במקום לדמות למידה של כמות (עוד שאלות), הגישה מדמה למידה איכותית, דומה לזו של בני אדם, בדומה להרהור חוזר של תלמיד על פתרון נתון. יתרון בולט של הגישה הוא שהיא לא משנה את אופן ההסקה בזמן ההפעלה (inference). כלומר, אין צורך לקרוא או לייצר את מקטע ההרהור בזמן ריצה, אך הוא כן משפיע על אופן החשיבה שנלמדת במהלך האימון. זהו שינוי עמוק בתפיסת "למידה על פתרונות" לעומת "למידה דרך הבנת עקרונות".

המאמר מדגים אמפירית שגישה זו לא רק משפרת את הדיוק בפתרון בעיות סטנדרטיות, אלא (וזה המרכיב החשוב) – משפרת בצורה יוצאת דופן ביצועים בתרחישים רפלקטיביים: תיקון טעויות, פתרון שאלות המשך, והסתמכות על משוב חיצוני. כמו כן, החידוש אינו מתנגש עם שיטות קיימות להעשרת דאטה (כגון Q-aug או A-aug) אלא משתלב עימן. השילוב בין reflective augmentation לבין הרחבות קיימות הביא לתוצאות הגבוהות ביותר בניסוי.

לסיכום, חידושו של המאמר נעוץ בשלושה מישורים:

מעבר ממודלים של "שינון פתרונות" למודלים של "הבנת עקרונות דרך הרהור".

חיקוי של למידה אנושית באימון של מודלים לשפה.

גישה חדשה שקל ליישם אותה על כל דאטה קיים מבלי לשנות את תהליך ההסקה.

מדובר בחידוש בעל פוטנציאל רחב שהתממש בעשרות מאמרים שיצאו בשנה האחרונה אחריו, במיוחד עבור פתרון שאלות מתמטיות על ידי LLMs, אך גם בתכנון סוכנים אינטראקטיביים הדורשים חשיבה גמישה ולא ליניארית.

https://arxiv.org/abs/2406.12050