המאמר היומי של מייק: 11.07.25

s1: Simple test-time scaling

סקירה קצרה של מאמר שיצא לפני חצי שנה שדי התפרסם בזמנו. האמת קצת איחרתי עם סקירותו אבל מרגיש שחובתי לסקורו כי יש כמה רעיונות מעניינים בנוגע ל test time compute או TTC. גישת TTC עלתה לכותרת לפני קצת יותר משנה ובעקרון היא אומרת על ידי שליטה וניהול של כמות הטוקנים שהמודל מגנרט במענה על שאלה עשויה להיות להוביל לשיפור בביצועי המודל במיוחד בשאלות שדורשות הנמקה (reasoning).

באופו מעניין (לפחות לדעתי) המודל קיבל שם s1, כאשר s בא ממילה small המתייחסת לגודל דאטהסט שהמודל עבור fine-tune עלי (1000 דוגמאות בלבד) וספרה אחת באה מ-o1 המודל הראשון של OpenAI שהשתמשו (בצורה מוצהרת) ב-test compute. המאמר מציע שני חידושים עיקריים: בניית דאטהסט והאלגוריתם ל-test compute עצמו.

הדאטהסט שהם בנו מורכב משאלות(והפתרון) מורכבות בדומיינים שונים כמו מתמטיקה, ביולוגיה, פיזיקה, כימיה וכדומה. כדי לבחור שאלות באמת מורכבות המחברים נתנו לשני מודלים של Qwen בגודל 32B ו- 7B. הפתרון של שני המודלים נבדק על ידי claude 3.5 ורק השאלות שנפתרו לא נכון על ידי שני המודלים נבחרו לדאטהסט. בשלב האחרון המחברים דאגו שכל דומיין יקבל ייצוג שווה פחות היותר כאשר בכל דומיין נבחרו שאלות עם פתרון הארוך ביותר (שכנראה משקף את קושי השאלה). בסוף המודל qwen 32B עבר SFT על הדאטהסט הזה.

החידוש השני הוא כאמר ה- test time compute בזמן האינפרנס. המחברים דאגו (על ידי הכנסה של טוקנים מסוימים כמו "wait" ו-״end of thinking" שתהליך החשיבה (כמות הטוקנים) שהמודל משקיע בפתרון לא יהיה לא ארוך מדי ולא קצר מדי. למשל אם הפתרון קצר מדי המחברים מוסיפים (דוחפים לתוך הטוקנים המגונרטים) את הטוקן "wait" וכאשר הפתרון ארוך מדי מכניסים את הטוקן ״end of thinking" והמודל נאלץ לתת את התשובה בהתבסס על שרשרת ההנמקה שכבר בנה.

המאמר מצא כי הארכה של חשיבת המודל בדרך כלל משפיע חיובית על דיוק הפתרון אבל אחר 4 הכנסות של "wait" הביצועים מפסיקים להשתפר. לפי המאמר אין יותר מדי השפעה לקיצור בכוח של שרשרת חשיבת המודל לפחות בגבולות אוך חלון ההקשר.

המאמר מראה ביצועים שווים פחות או יותר עם המודלים שאומנו על דאטהסטים גדולים הרבה יותר שמדגיש חשיבות איכות הדאטה לאימון מודלים. בנוסף גם test-compute בזמן האינפרנס בטח תרם לביצועי המודל.

https:/axiv.org/abs/2501.19393