⚡️🚀המאמר היומי של מייק 03.09.24: ⚡️🚀

Smaller, Weaker, Yet Better: Training LLM Reasoners via Compute-Optimal Sampling

אחת הדרכים הדי מפתיעות לשיפור יכולות reasoning של מודלי שפה היא שיפור עצמי או self-improvement. בגדול עבור דאטהסט של שאלות ותשובות אנו מבקשים ממודל שפה לענות על התשובה ולספק הסבר. לאחר מכן מפלטרים את השרשראות reasoning שלא התכנסו לתשובה הרצויה. לאחר הפלטור מבצעים פיינטיון של המודל על הדאטהסט המפולטר. וכאמור באופן די מפתיע (לפחות אותי) הדבר אכן מוביל לשיפור יכולות reasoning של מודל שפה.

ואם יש בידינו מודל יותר חזק אז ניתן לבנות את הדאטהסט הזה באמצעותו ולעשות את הפיינטיון על הדאטה הנוצר באמצעותו בצורה דומה.

אולם המאמר שואל שאלה די מעניינת: מה עדיף (מבחינת הביצועים), ליצור יחסית מעט דאטה עם מודל גדול וחזק או ליצור יחסית הרבה דאטה עם מודל קטן וחלש יותר. הרי יצירת דאטה עם מודל חזק היא יקרה יותר (מבחינת כמות ה-FLOPS הכוללת הנדרשת לכך) אבל מצד שני הדאטה שהוא יוצר הוא יותר איכותי.

המחברים מציעים לבצע את ההשוואה של ״תפוזים לתפוזים״ - כלומר לקחת את הדאטה הנוצר עם מודל חזק ומודל חזק תחת אותו תקציב של FLOPS ולהשוות מה מהם מוביל לביצועים טובים יותר של המודל שעובר פיינטיון על הדאטה הזה.

ויש תוצאות די מעניינות במאמר..

.

https://arxiv.org/pdf/2408.16737