🚀המאמר היומי של מייק 17.09.24: ⚡️🚀  
STaR: Self-Taught Reasoner Bootstrapping Reasoning With Reasoning

אני ממשיך לחפור במאמרי שאולי עיצבו את הנתיב הובילו ל-o1 של openai. הפעם נברתי כה עמוק שהגעתי למאמר שיצא לפני שנתיים וחצי (בדיפ היום זה כמו 100 שנה במתמטיקה). שימו לב שהמאמר יצא עוד לפני chatgpt. המאמר הזה מציע שיטה לשיפור יכולת reasoning של מודל שפה כאשר בידנו יש דאטהסט גדול של שאלות ותשובות D ודאטהסט קטן D\_R הרבה יותר (המאמר מדבר על 10 דוגמאות בלבד) המכיל בנוסף גם את שרשרת ה-reasoning.

כאשר אני מדבר על שיפור איכות ה-reasoning אני בעצם מתכוון לפיינטיון של המודל במטרה לקבל מודל חזק יותר ב-reasoning. המחברים מציעים אלגוריתם המורכב משני שלבים עיקריים. בשלב הראשון מזינים את הבאץ' של שאלות למודל שפה כאשר בנוסף לשאלות הפרומפט מכיל את דוגמאות לשרשראות ה-reasoning m מ- D\_R. המודל מתבקש לבנות שרשרת reasoning לכל השאלות מבאץ' (לא מ-D\_R) ולהגיע לתשובה הסופית.

את שרשראות ה-reasoning לשאלות שהצליחו להגיע לתשובה נכונה מוסיפים לסט שנקרא לו D\_N. לשאלות שהמודל לא הצליח להגיע לתשובה סופית נכונה אנחנו מוסיפים רמז (במאמר זה נקרא rationalization) שעוזר למודל לבנות את שרשרת ה-reasoning. השאלות שהצליחו להגיע לתשובה הנכונה אחרי הרמז גם נוספים ל D\_N. לאחר מכן מבצעים איטרציה אחת של שיטת מורד הגרדיאנט נבחרת על D\_N ומעדכנים את משקלי המודל. חוזרים על השלבים האלו עד שהלוס מתייצב.

זהו זה, שיטה אינטואיטיבית ופשוטה שקיבלה כמה מאמרי השמך די כבדים שבתקווה אסקור אותם גם כן

https://arxiv.org/pdf/2203.14465