⚡️🚀המאמר היומי של מייק 10.08.24: ⚡️🚀

Synthesizing Text-to-SQL Data fromWeak and Strong LLMs

הסקירה של היום הולכת להיות די קצרה וקלילה. המאמר מציג שיטה די אינטואיטיבית לאמן מודל שפה קטן לבצע משימה מסוימת. במקרה שלנו המשימה היא גנרוט של שאילתת SQL לפי תיאורה הטקסטואלי ומבנה (schema) של הטבלה. מודלי שפה קטנים עלולים להסתבך עם המשימה הזו בטח במקרים שהשאילתה הנדרשת אינה טריוויאלית.

המאמר מציע תהליך דו שלבי של אימון מודל קטן למשימה זו. בשלב הראשון יוצרים דאטהסט עבור המשימה הזו באמצעות מודלי שפה גדולים וחזקים וכמה דאטהסטים רלוונטיים. עושים דברים רגילים, הנדסת פרומפטים קלה וכאלו. לאחר מכן עושים למודל הקטן פיין טיון על הדאטהסט הזה.

בשלב השני עושים למודל השפה הקטן Direct Policy Optimization או DPO שראינו אותו כשלב אימון מודלי יסוד (foundational). היתרון של שיטה זו היא בכך שהיא לא דורשת אימון של מודל reward. בשביל אימון מודל כזה אנו צריכים דוגמאות טובות ודוגמאות לא טובות. דוגמאות טובות יש לנו מהשלב הראשון.

בשביל לבנות את הדוגמאות הרעות לוקחים את המודל הקטן המתקבל על השלב הראשון כדי לגנרט שאילתת SQL לתיאור טקסטואלי נתון. לאחר מכן מריצים את השאילתה כדי לוודא האם התוצאה המתקבלת נכונה. אם היא לא נכונה קיבלנו דוגמא שלילית. ככה בונים דאטהסט של דוגמאות חיוביות ושליליות ומה ש נותר לעשות הוא PPO.

https://arxiv.org/abs/2408.03256