המאמר היומי של מייק - 10.12.24:  
LLM2LLM: Boosting LLMs with Novel Iterative Data Enhancement

#### 1. מבוא ומוטיבציה

המאמר מציג את LLM2LLM, מסגרת חדשנית לשיפור ביצועי LLMs במצבים של מחסור בדאטה. בעוד שאימון נוסף של מודלים כאלה דורש בדרך כלל דאטה מתויג רב, מה שדורש עבודה ידנית מרובה, LLM2LLM מציע אסטרטגיית העשרת דאטה איטרטיבית המבוססת על פרדיגמת מורה-תלמיד(student-teacher) כדי לשפר את הדאטה בעייתיות (שהמודל הקטן, תלמיד, מתקשה להתמודד איתם) באופן דינמי

2. מתודולוגיה

LLM2LLM מורכב מ-3 שלבים עיקריים:

אימון מודל התלמיד: מודל התלמיד מאומן על כמות דאטה קטנה.

זיהוי שגיאות: הביצועים נמדדים על נתוני האימון באמצעות המודל הגדול (מורה), ודוגמאות שבהן המודל הקטן שוגה מזוהות.

העשרת דאטה ממוקדת: מודל המורה מייצר דוגמאות סינתטיות חדשות בתור אוגמנטציות שונות של הדוגמאות בהם מודל התלמיד טועה. דוגמאות אלו משתלבות מחדש במערכת לצורך איטרציות אימון נוספות.

מאפיינים מרכזיים:

העשרה איטרטיבית: הדאטהסט לאימון מודל התלמיד משתפרות לאורך מספר סבבים במקום להיווצר מראש.

מיקוד בטעויות: הדגש הוא על דוגמאות מאתגרות המדגישות את חולשות המודל הקטן.

המחברים מציינים כי מודל המורה אינו חייב להיות חזק יותר, אלא רק להפיק דוגמאות קונספטואליות דומות לטעיות של המודל הקטן.

#### 3. תוצאות

המסגרת הוכיחה שיפורים משמעותיים במדדים במצבי מחסור בדאטה תוך שהיא מתעלה על שיטות העשרה אחרות. דוגמאות לביצועים:

GSM8K (הסקה מתמטית): שיפור של 24.2% בדיוק.

CaseHOLD (הסקה משפטית): שיפור של 32.6%.

SNIPS (זיהוי כוונות): שיפור של 32.0%.

TREC (סיווג שאלות): שיפור של 52.6%.

SST-2 (ניתוח רגשות): שיפור של 39.8%.

#### 4.סיכום

LLM2LLM מציעה מסגרת להעשרת הדאטה באימון LLMs במצבים של מחסור בדאטה. על ידי התמקדות איטרטיבית בדוגמאות מאתגרות ושימוש בשיתוף פעולה בין מורה לתלמיד, היא משיגה שיפורי ביצועים משמעותיים. שיטה זו מסמנת כיוון מבטיח לשיפור היעילות והשימושיות של מודלים לשוניים בסביבות מוגבלות משאבים.

https://arxiv.org/pdf/2403.15042