⚡️🚀המאמר היומי של מייק -12.11.24: ⚡️🚀

OccamLLM: Fast and Exact Language Model Arithmetic in a Single Step

זהו מאמר שממש אהבתי, אהבתי גם את הרעיון וגם כתוב בצורה מאוד ברורה. למה כה אהבתי את הרעיון? אני כבר זמן מה טוען שבמקום להשקיע מאמצים גדולים באימון מודלי שפה לפתור בעיות מתמטיות יחסית מורכבות (שלדעתי מאוד קשה כי הם לא ״בנויים״ לזה באופן טבעי) כדאי להשתמש בכלים חיצוניים ייעודיים לכך (למשל כלים סימבוליים). מטרה של מודלי שפה במקרה הזה היא לזהות מתי הקלט שמוזן אליו (הפרומפט) מצריך פתרון בעיה מתמטית, ״לתרגם״ את הבעיה לשפה של הכלי הייעודי הזה, להעביר את הבעיה המתורגת לשפתו אליו לפתרון ולפענח את הפלט שלו.

וזה בדיוק מה שהמאמר הזה עושה. המחברים לקחו מודל שפה ופתחו מודל נפרד לפתרון בעיות מתמטיות. למעשה המודל לפתרון בעיות מתמטיות שפותח במאמר הוא גרף חישובי דינמי שכל צומת בו היא פונקציה או פעולה מתמטית (נדיג סימן + ו- \*, או cos ו-exp). יש גם צמתים למשתני קלט השונים כדי שהמודל יוכל לחשב פונקציות על כמה משתנים (multivariate). למעשה גרף כזה הוא DAG או בשמו המלא Directed Acyclic Graph ומאמנים אותו לבחור את ״נתיב החישוב״ בו (״מסלול הצמתים״) בהינתן הייצוגים (אמבדינגס של הטוקנים) המוחשבים על ידי מודל שפה (ד״א מודל שפה לא מאומן ונותר קבוע לכל אורך אימון המודל).

המחברים מאמנים שני מודלים: הראשון מזהה האם יש צורך בהפעלת המודל לחישובים מתמטיים לכל טוקן בהינתן ההקשר (כלומר כל הטוקנים לפניו). המודל השני מאומן לבנות נתיב חישובי בגרף החישובי שתיארתי בפסקה הקודמת. את שני המודלים האלו מאמנים בנפרד.

מעניין כל שכבה של רשת ה-DAG הזה מורכבת משני חלקים: בחלק בראשון יש לנו צמתי החלטה: כל צומת כזה הוא וקטור ״המחבר״ אותו לצמתים פונקציונליים שכל אחד מהם הוא בעצם פעולה או פונקציה מתמטית (מקבוצת פעולות ופונקציות שבחרנו). הוקטור הזה הוא למעשה סופטמקס שממנו נדגם לאיזה צומת פונקציונלי/פעולה נחבר אותו. כל צומת פונקציונלי שנבחר מחובר עם כל צמתי ההחלטה מהשכבה הבאה ואליהם מועבר הייצוג משכבת ההחלטה הקודמת יחד עם ייצוג הפעולה (כנראה האם נבחרה או לא). כך נבנה גרף חישובי מייצוגי הטוקנים המחושבים על ידי מודל שפה (הם מחוברים לשכבת ההחלטה הראשון במודל החישובי). ד״א כל פעולה וכל פונקציית בסיס בגרף משוכפלת בכמה צמתית כדי להקנות למודל יכולת לקרב פונקציות מורכבות יותר.

מכיוון שאנו דוגמים את הגרף החישובי כל פעם מחדש עבור כל פלט של מודל השפה, לא ניתן לאמן אותו בקלות על שיטות קלאסיות של למידת מכונה (supervised learning). המחברים בחרו בשיטה קלאסית מעולם למידה עם חיזוקים (RL) הנקראת reinforce כאשר פונקציית reward היא עד כמה התשובה המחושבת באמצעות הגרף החישובה קרובה לתשובה ground truth. דרך אגב ניתן לייצג רוב הפונקציות עם עם יותר מאחד נתיבי חישובי.

מאמר די נחמד אבל כתוב לא מאוד ברור (או שהיה חסר לי קצת רקע)...

https://arxiv.org/abs/2406.06576