⚡🚀 המאמר היומי של מייק 31.05.24: ⚡🚀

Transformers Can Do Arithmetic with the Right Embeddings

אנו יודעים שמודלי שפה גדולים לא מצטיינים בלחשב ביטויים מתמטיים בטח כאלו המכילים מספרים עם הרבה ספרות. גם אם מאמנים אותם על מיליוני דוגמאות עדיין מסתבכים להכליל אותם למספרים גדולים. המאמר מציע להוסיף positional encoding למספרים שמטרתם לספק למודל שפה מרחק של כל ספרה מתחילת המספר. וזה עובד לא רע.

רפו: https://github.com/mcleish7/arithmetic

מאמר: https://arxiv.org/abs/2405.17399