⚡🚀המאמר היומי של מייק 02.06.24: ⚡🚀

LLaMA-NAS: Efficient Neural Architecture Search for Large Language Models

פעם הנושא של Neural Architecture Search או NAS בקצרה שעסק בחיפוש לאחר ארכיטקטורה אופטימלית של רשת נוירונים עבור משימה/משימות/דומיין היה די פופולרי אך בשנים האחרונות התחום נמצא בדעיכה. אני שמח שנתקלתי במאמר הזה שמנסה לפתח NAS עבור מודלי שפה. אני זוכר מאמרים די מגניבים שמשתמשים בשיטות RL די מגניבות לכך. אולי בעתיד NAS תהפוך למתחרה רציניות של שיטות פרונינג וקוונטיזציה.

ֿ

מאמר: https://arxiv.org/abs/2405.18377

טלגרם: https://t.me/MathyAIwithMike/69