⚡️🚀המאמר היומי של מייק 27.06.24:⚡️🚀

Probing the Decision Boundaries of In-context Learning in Large Language Models

המאמר הזה מדגים בפעם מי יודע מה שיש משימות שמודלי שפה מתקשים בהם מאוד אבל אם נסביר לו את המשימה ב״שפתו״ הוא די מצליח להסתדר איתה. הפעם המשימה היא סיווג מולטיקלאס - כלומר אנו מספקים למודל כמה זוגות של וקטורי x והלייבל שלו y. הווקטורים ניתנים להפרדה בצורה לינארית על ידי ישר מסוים כלומר נמצאים בשני צידיו (של הישר). זה ה-context שלנו. לאחר מכן המודל מקבל נקודה x ומתבקש לחזות את הלייבל שלו y.

המחברים ניסו להבין עד כמה המודל מצליח לזהות את הלייבלים של הנקודות הנמצאות בין המינימום למקסימום של נקודות x שנתנו לא ב-context. בשביל כך הם חקרו את את התפלגות הטוקן הבא אחרי השאילתה עבור כל הקטגוריות עבור ערכי x שונים. באופן לא מפתיע המודל לא למד להפריד את הנקודות על ידי הישר אלא התכנס לקו הפרדה די פרוע ורחוק מהישר המפריד. הגדלת מספר הנקודות ב-context לא עזר ובנוסף קו ההפרדה היה רגיש לסדר הנקודות וגם לסימון של הלייבלים השונים.

אז המחברים גילו שפיינטיון של המודל על משימות דומות די עוזר למודל להתכנס לפתרון הנכון (גם לא מפתיע). ויש עוד כמה דרכים לגרום למודל לפתור את המשימה הפשוטה הזו.

אם אתם שואלים אותי למשימות כאלו יש לכם רגרסיה לוגיסטית….

https://arxiv.org/abs/2406.11233