⚡️🚀המאמר היומי של מייק 24.07.24: ⚡️🚀

The Empirical Impact of Neural Parameter Symmetries, or Lack Thereof

הסקירה היום תהיה קצרה וקלילה לעומת הסקירות האחרונות על מודלי דיפוזיה למיניהם. המאמר של היום חוקר סימטריות ברשתות נוירונים עמוקות. ניתן לראות די בקלות כי קיימות לא מעט פרמוטציות של המטרצות המשקלים בשכבות השונות של רשת שלמעשה לא משנות את המודל. כלומר אם תפעילו את המודל אחרי פרמוטציה על כל קלט תקבלו את אותה התוצאה כמו עם המודל המקורי.

האם הסימטריות האלו מביאות לנו משהו טוב? בכלל לא בטול - לי זה נראה (למרות שאני לא מומחה גדול בתחום) כמו סוג של יתירות של יש במודלים שבלעדיה אולי ניתן היה להגיע למודלים קטנים יותר למשל. המאמר בוחן מה קורה במודל עם אנו מפרים את הסימטריה שיש במודל. אחת הדרכים להרוס את הסימטריה היא לקבע משקלות (לערכים אקראיים אך קבועים) במקומות שנבחרו באקראי במטריצות משקלים של הרשת. הדרך השניה היא להפעיל פונקציה אקטיבציה רק על המשקלים מסוימים.

המאמר חוקר איזה אפקטים מתרחשים אחרי שהורסים את הסימטריה במודל ומגלה כמה דברים די מעניינים….

https://arxiv.org/pdf/2405.20231