המאמר היומי של מייק - 21.03.25  
LLMs learn governing principles of dynamical systems, revealing an in-context neural scaling law

המאמר משך את תשומת ליבי כי מופיעים בשמו מודלי שפה ומערכות דינמיות שאני מחבב מהזמנים העליזים של ממבה (state-space models). המאמר טוען שמודלי שפה מפגינים ביצועים טובים בהבנת מערכות דינמיות ממגוון סוגים כולל מערכות סטוכסטיות, כאוטיות, רציפות וכדומה. וכל זה קורה ללא שום טיוב (fine-tune) - כלומר קצת הנדסת פרומפטים ומודל השפה שלכם מבין במערכות דינמיות.

לאחרונה היו לא מעט מאמרים שניסו לפצח ״מערכות דינמיות״ הניתנות על ידי דגימות שלהם עם LLMs דרך יצירת דגימות חדשות ממנו (מהמערכת הדינמית) באמצעות LLMs. ההיגיון כאן די פשוט - אם מודל שפה יודע לגנרט מהתפלגות המושרית על ידי מערכת דינמית, אז כנראה הוא מבין אותה.

המחברים לקחו גישה אחרת יותר ישרה - הם הראה שניתן ממש ליצור התפלגות של מערכת דינמית באמצעות LLM כאשר המערכת היא מרקובית. כלומר אם התפלגות דגימה הבאה בזמן t+1 תלויה רק במצב המערכת בזמן t ולא בעבר. עבור מערכת דיסקרטית התפלגות זאת נתונה על ידי מטריצה של הסתברויות מותנות המכילה את ההסתברויות של מצב { x\_{t+1} בהינתן מצב x\_t בזמן עבור כל הערכים האפשריים שלהם. עבור מערכות רציפות ניתן לבנות מטריצה כזו על ידי דיסקרטיזציה של הערכים של מצבי המערכות.

המאמר מראה שמודלי שפה מצליחים לבנות את מטריצות מעברים בצורה לא רעה במיוחד במצבים שיש יחסית מעט מצבים אפשריים. המרחק בין ההתפלגות החזויה על ידי מודל שפה לבין התפלגות ground truth נמדדה במאמר עם מרחק Bhattacharyya שנתקלתי בו רק בפעם השניה במאמרי deep learning. אציין שהמאמר מציג תוצאות טובות גם עבור מרחקים (divergence) אחרים כמו JSD ו- KL. המחברים מציעים דרך טריקית לבנות את המטריצה הזו עם LLM - מי שרוצה לצלול לעומק, תראו פרק שנקרא Hierarchy-PDF algorithm.

וזהו זה, היום זה היה קצר…

https://arxiv.org/abs/2402.00795