# Review 180: [Short] Learning From Mistakes Makes LLM Better Reasoner

**Paper: https://arxiv.org/abs/2310.20689v4**

https://huggingface.co/papers/2310.20689

טוב, שוב בשבת אנחנו במאמר קליל וקצר על מודלי שפה גדולים. מתברר שמודלי שפה בדומה לבני אנוש מסוגלים ללמוד מהטעויות שלהם. כלומר אם אנו נכייל מודל שפה על הדאטה המכיל תיקונים של טעויות חשיבה המודל ילמד יותר טוב. כלומר המאמר לוקח מספר דוגמאות ומבקש ממודל שפה לבנות שרשראות חשיבה (Chain-of-Thought- CoT) עבורן.   
  
לאחר מכן ומפעילים על שרשראות חשיבה אלו מודל שפה חזק (כמו GPT4) ומבקשים ממנו לחפש שגיאות. המודל מתבקש לאתר שלבים המכילים שגיאות ולתקן אותם עד הפתרון. המחברים מצאו ש-GPT4 די טוב במשימה הזו והצליח למצוא שלבים בעייתים ברוב שרשראות החשיבה.

אז מה שהמחברים עשו הם לקחו דאטהסט של (שאלות, ותשובות), הפעילו את GPT4 עליו כדי לבנות שרשרת חשיבה. הם יצרו כמה שרשראות חשיבה לכל שאלה ופלטרו שרשראות על תשובות לא נכונות. לאחר מכן הם כיילו מודל שפה(פחות חזק מ-GPT4) על הדאטה הזה. בשלב הבא הם הפעילו את GPT4 כדי לתקן שגיאות בשרשראות הלא נכונות וטייבו (finetune) את המודל עליו. כתוצאה מכך הביצועים של המודל במשימה השתפרו. זהו זה – קליל וקצר לשבת.