# Review 173: [Short] Teaching Language Models to Self-Improve through Interactive Demonstrations

**Paper: https://arxiv.org/abs/1909.04157v1**

https://huggingface.co/papers/2310.13522

ממודלי דיפוזיה שסקרנו אתמול עוברים לאייטם פופולרי אפילו אפילו מהם כלומר מודלי שפה ענקיים (LLMs). המאמר שנסקור היום מציע שיטה לאימון LLMs קטנים יחסית (מיליארדי פרמטרים בודדים) לפתרון בעיות מתמטיות מורכבות (נניח כאלו שמכילות הרבה פעולות).

המאמר מציין שמודלי קטנים יחסית מתקשים לפתור בעיות בעזרת reasoning אם מפעילים אותו (המודל) בצורה של few-shot, כלומר מספקים לו כמה דוגמאות עם פתרון מלא. בגדול המאמר מציע לאמן (מכייל) מודל שפה קטן L על הטעויות שלו. עבור בעיה נתונה מפעילים מודל L כדי ליצור שרשרת צעדי חשיבה לפתרון בעיה זו. לאחרי מכן מפעילים מודל יותר חזק (נגיד codex) לפתרון בעיה זו ומשווים את שרשרת החשיבה של שניהם.

במקום הראשון שהם שונים מחליפים את המשוב של המודל החלש בזה של המודל החזק. לאחר מכן מפעילים מודל חזק שוב פעם כדי לתקן את שרשרת החשיבה של המודל החלש מהמקום הזה. לאחר מכן מחלקים את הדאטהסט הזה (יש בו פתרונות זהב ground-truth, פתרונות נכונים של המודל החלש, והפתרונות המתוקנים על ידי המודל החזק). אז מחלקים את הפתרונות האלו לפי התוצאה הסופית (נכונה או לא נכונה).

את הפתרונות הנכונים מחלקים לשלישיות של (תוצאה של שלב i, המשוב והתוצאה של השלב החדש). אלו שמסתיימים בפתרון האחרון מחלקים לזוגות (שלב i, משוב). בסוף מאמנים מודל קטן על הדאהטסט הזה תוך משקול שונה לשלישיות והזוגות מהשלב הקודם. מטרת האימון היא חיזוי הטוקן הבא כמו שמקובל באימון מוקדם של מודלי שפה. ככה מצליחים לשפר את הביצועים של המודל הקטן במשימות מורכבות של reasoning.