# Review 142: LARGE LANGUAGE MODELS AS OPTIMIZERS, 09.09.2023 https://huggingface.co/papers/2309.03409

**Paper: https://arxiv.org/abs/2311.15249v1**

מודלי שפה נמצאים היום כמעט בכל מקום: הם עוזרים לנו לכתוב תוכן, לבצע מגוון משימות הקשורות לשפה טבעית כמו תרגום, ניתוח סנטימנט, אנו מנהלים איתם דיאלוגים מעניינים להנאתנו. אבל האם הם מסוגלים לפתור בעיות אופטימיזציה כמו רגרסיה לינארית או בעיית איש המכירות המטייל?  
  
היום ב-#shorthebrewpapereviews אנו סוקרים קצרות מאמר שמראה שמודלי שפה כן מסוגלים לפתור בעיות אופטימיזציה הנ״ל, כמובן אם מדברים איתם יפה (כלומר מהנדסים פרומפטים בצורה מתאימה) אז הם מצליחים לפתור בעיות רגרסיה לינאריים (במימד אחד, כלומר למצוא שני מקדמים של הישר) וגם בעיית איש המכירות המטייל (למצוא מסלול הקצר ביותר המבקר בסט נקודת – כאן על המישור).   
  
במקרה של רגרסיה בוחרים ערכים של שני המקדמים (w,b) של רגרסיה דוגמים 50 נקודות x ומחשבים בהם את ערך הפונקציה בתוספת רעש גאוסי. מתחילים כמה זוגות אקראיים של w ו- b, מחשבים את השגיאות על הדאטהסט עבור הערכים הנבחרים של w ו- b שנדגמו. נותנים את השגיאות האלו למודל שפה ומבקשים ממנו לתת ערכים של w ו- b שממזערים את ההפרש הזה.   
  
המודל מנחש ואז מספקים לו כמה זוגות של ערכי w ו- b המוצלחים ביותר. והמודל מצליח די מהר להגיע די קרוב לתשובה הנכונה. דבר דומה עושים לבעיית איש המכירות המטייל וגם שם מודל שפה די מצליח. אציין לי שלא ברור לי איך מעבירים את הדאטהסט למודל. אחר כך המחברים ביצועי אופטימיזציה של הפרומפט (נקרא meta-prompt) במטרה למזער את השגיאה על הטסט סט. מטה-פרומפט זה מורכב משני דברים:

הפרומפטים הקודמים שנוסו והדיוקים של הפתרונות (שערוכי המקדמים) שהמודל סיפק באיטרציות הקודמות.

תיאור הבעיה יחד עם הדגימות מהטריין סט (ממש בחירת מיניבאטץ')

התוצאות די נחמדות…