# Review 164: [Short] Table-GPT: Table-tuned GPT for Diverse Table Tasks

**Paper: https://arxiv.org/abs/2310.09263v1**

https://huggingface.co/papers/2310.09263

אתם בטח מודעים ליכולות המטורפות של מודלי שפה אבל הם עדיין מתקשים להסתדר עם דאטה טבלאי. המאמר שנסקור היום ב-#shorthebrewpapereviews מציע שיטת טיוב(fine-tuning) של מודלי שפה שבאה להקנות להם יכולת לעבוד עם טבלאות.   
  
קודם כל בואו נבין למה מודלי שפה בעצם מתקשים לשחזר את הביצועים החזקים שלהם בדאטה טבלאי. הסיבה נעוצה בשינויים בין המאפיינים המהותיים של דאטה טבלאי ושפה טבעית. הטקסט הוא חד כיווני (או משמאל לימין או מימין לשמאל כמו עברית) ולעומת זאת לטבלאות מבנה דו-ממדי. דאטה טקסטואלי לא אינווריאנטי לפרמוטציות לעומת רוב הטבלאות שפרמוטציה של עמודות או של שורות אינה משפיעה על תכונות הטבלה. המחברים מציעים לכייל מודל שפה על המשימות שהן אינהרנטיות לטבלאות שהן כמובן מאוד שונות מהמשימות שאנו רואים בעיבוד שפה טבעית.   
  
למשל אחת המשימות שמודל שפה מכויל עליהן היא זיהוי מקומות בטבלה שבהם יש דאטה חסר. משימה אחרת (טיפה יותר מורכבת) היא לאתר שורות בשתי טבלאות המייצגות את אותו ה"ישות״ (entity). עוד משימות טבלאיות היא השלמה ערכים חסרים בטבלה, הפיכה של שאלה מילולית ששאילתה עבור הטבלה ותמצות של תוכן הטבלה. יש כמעט 20 משימות שונות שעליהן מכיילים מודל שפה והמודל המכויל הנושא שם הלא מפתיע TableGPT מציג ביצועים די טובים.