# Review 154: Context is Environment, 26.09.2023

**Paper: https://arxiv.org/abs/2309.09888v2**

https://arxiv.org/abs/2309.09888

סקירה זו נכתבה על ידי עדן יבין

מודל של רכב אוטונומי צריך לדעת להתמודד עם המון מצבי עולם אשר לא ראה בזמן תהליך האימון. כיצד הוא עושה זאת? מזעור הטעות על דוגמאות כאלו הינו תחום שלם הנקרא Domain Generalization. האם מודלי שפה יוכלו לעזור לתחום זה ולהראות שיפור על פני המצב הקיים?

נגלה היום ב- #shorthebrewpapereviews. המאמר נקרא Context is Environments ובמסגרתו החוקרים מנסים להראות שמה שקוראים לו ״הסביבה״ בתחום ה-DG מקביל מאוד ל-Context בתחום של מודלי שפה. שיטות קיימות בתחום ה-DG מנסות להשתמש בדוגמאות העבר שנוצרו מאינטראקציה עם הסביבה כדי לחזות את התוצאה של הדגימה הנוכחית.   
  
אך האם לא כך גם אצל מודלי שפה? הרי הם משתמשים בטוקנים הקודמים כדי לחזות את הטוקן הנוכחי. יותר מכך, עם השימוש ההולך וגובר במודלי שפה גילו את היכולת שלהם ללמוד in-context באמצעות טכניקות כגון few-shot. נוכל להשתמש בכך בשביל לשפר את יכולת ההלכה של מודלים אלו על דוגמאות אשר לא ראו. השיטה של החוקרים נקראת ICRM, ובקצרה מנסה להשתמש בקונטקסט כסביבה כדי להקטין את הסיכוי לטעויות על דוגמאות אשר לא נראו ולא דומות למה שהיה באימון המודל.

כאשר מודל השפה h מנסה לשערך את (P (Y| X,C על ידי שימוש בפונקציית הפסד של binary cross-entropy loss. השערוך של (P(Y|X,C) הינו בשביל לשערך את הסיכון של טעות בחיזוי בהינתן הדוגמא הנוכחית והסביבה או הקונטקסט.

החוקרים מראים ששימוש פשוט זה מביא לתוצאות טובות יותר מהשיטות הקודמות בניסויים הכוללים יכולת הכללה על דוגמאות חדשות שלא נראו בסט האימון. למי שירצה להתעמק יותר, המאמר מראה עוד המון נקודות קריטיות וחשובות בשימוש של מודלי שפה בשביל לחשב סיכון של דוגמאות חדשות ובנוסף נותן עוד תאוריה מעניינת על התחום.