המאמר היומי של מייק - 02.01.25:  
Loss of plasticity in deep continual learning

היום סוקרים קצרות מאמר די קליל מ-nature.

מבוא:

שיטות למידה עמוקה סטנדרטיות מציגות ירידה הדרגתית ביכולתן ללמוד משימות חדשות בצורה מתמשכת(״מוסיפים״ למודל משימה בצורה הדרגתית). בניגוד לשכחה קטסטרופלית(catastrophic forgetting), שבה ידע קודם אובד, אובדן פלסטיות מגביל את יכולת הרשת ללמוד משימות חדשות ביעילות.

ניסויים מקיפים על דאטהסטים כמו ImageNet ו-CIFAR-100, כמו גם תרחישי למידה עם חיזוקים (Reinforcement Learning), חשפו שהנוירונים הופכים רדומים (לא משתנות בכל הדוגמאות) או מתמחות יתר על המידה על משימה ספציפית, מה שמפחית את יכולתן להסתגל לדאטה חדש. לאורך זמן, רשתות החוות למידה מתמשכת מתפקדות לא טוב יותר ממודלים רדודים (לינאריים), מה שמדגיש מגבלה בסיסית של שיטות מבוססות מורד הגרדיאנט(gradient descent) ללמידה מתמשכת (ואנו מאמנים מודלים עם GD היום)....

מורד הגרדיאנט ללמידה מתמשכת:

שיטות למידה מתמשכת מנסות להתמודד עם אובדן פלסטיות על ידי אתחול מחדש של נוירונים רדומים (כאלו שלא ״נדלקים כמעט אף פעם) ואימונם מחדש עם מורד הגרדיאנט. ככה גישה זו מנסה ״ליצור״ על נוירונים שילמדו משימה חדשה בלי להינעל על למשימות מסוימות, וזה שמאפשר לה ללמוד משימות חדשות ללא הידרדרות משמעותית בביצועים.

בניגוד לשיטות קונבנציונליות המסתמכות אך ורק על מורד הגרדיאנט, GD ללמידה מתמשכת מתאפיין בעדכון הדרגתי סטים שונים של משקלי המודל בדומה למה שקורה במערכות למידה ביולוגיות.

שיטות אימון נוספות:

כאמור אובדן פלסטיות קשור לאופטימיזציית יתר (לטענת המאמר) של משקולות והופעת נוירונים רדומים ברשת. נוירונים אלו אלה או מפסיקים לתרום ללמידה (עבור אקטיבציית ReLU) או נכנסות למצב רוויה(מגיעות ל 0 או 1 עבור סיגמואיד). טכניקות כמו רגולריזציית L2 מפחיתות את גדילת משקלי המודל ושומרות על ״פלסטיות״ (גמישות למשימות חדשות) במידה מסוימת. למשל שיטת Shrink and Perturb, המשלב רגולריזציה עם שינויים אקראיים קטנים במשקולות, מפחית את תופעת הנוירונים הרדומים וכך מגדיל את יכולת למידה של המודל.

אתגרי למידה מתמשכת ב-RL

למידה מתמשכת היא חיונית גם ל-RL אפילו יותר מאשר בבלמידה מפוקחת. לא רק שהסביבה יכולה להשתנות, אלא גם ההתנהגות של הסוכן הלומד יכולה להשתנות, ובכך להשפיע על המידע שהוא מקבל גם אם הסביבה נשארת קבועה. מסיבה זו, הצורך בלמידה מתמשכת הוא לעתים קרובות יותר ברור בלמידה עם חיזוקים, וRL היא סביבה חשובה להדגמת הנטייה של למידה עמוקה לאובדן פלסטיות. והמאמר בוחן שימוש בשיטות שדנו בהם קודם למשימות של RL יחד עם PPO, האלגוריתם המפורסם לאופטימיזציה ב-RL   
  
https://doi.org/10.1038/s41586-024-07711-7