המאמר היומי של מייק: 07.06.25  
Rate-In: Information-Driven Adaptive Dropout Rates for Improved Inference-Time Uncertainty Estimation

היום אני סוקר מאמר מיוחד בכמה רבדים. הרובד הראשון אחד ממחבריו של מאמר זה הוא לא אחר אלא יאן לקון, אחד האבות של למידה עמוקה. הרובד השני מכיל את החוקר הישראלי הידוע (אך לא מספיק) רביד זיו שוורץ שהוא גם פרופסור באוניברסיטת ניו יורק. הרובד השלישי הוא נושא המאמר והוא שערוך אי ודאות עבור חיזוים של רשתות נוירונים - נושא שמאוד מעניין אותי אך לא מעט זמן לא סקרתי כזה.

איך ניתן לשערך אי הוודאות של החיזויים של רשת נוירונים? יש כמה משפחות של שיטות המוזכרות במאמר:

רשתות נוירונים בייסיאניות מגדירות התפלגויות הסתברותיות על משקלי הרשת, מה שמאפשר למדל אי-ודאות דרך ההתפלגות הפוסטריורית. עם זאת, הן כבדות חישובית וקשה להרחיב אותן.  
שיטות אנסמבל: מאמנות מספר מודלים ומאגדות את התחזיות שלהם. מסוגלות למדל גם אי ודאות אפיסטמית וגם אליאטורית, אך דורשות משאבים חישוביים רבים.  
אוגמנטציה של דאטה בזמן טסט (Test-Time Augmentation): מוסיפות שיבושים לקלט (כמו סיבוב או טשטוש) כדי להעריך את התפלגות התחזיות. יעיל בעיקר כשיש ידע מוקדם על מבנה הנתונים.  
הזרקת רעש למודל: מוסיפים רעש נשלט (למשל גאוסי) למשקלים או לפעולות כדי לבחון רגישות מעבר לשינויים בקלט.  
שיטות מונטה קרלו (MC): משתמשות בדגימות אקראיות כדי לאמוד אי ודאות. למשל, MC Dropout מפעיל דרופאוט(dropout) גם בזמן טסט כדי לדגום את מרחב משקלי הרשת. יש לא מעט שיטות נוספות מבוססת MC לשערוך אי ודאות ברשתות.

אבל איך ניתן לשערך את הוודאות? אחת הדרכים היא להשתמש בגישות מתורת המידע (information theory) לניתוח של זרימה המידע בתוך הרשת ומידת ״פגיעתה״ מהשיטות המוזכרות מעלה (למשל MC Dropout). בגדול מאוד ככל שזרימת המידע נפגעת יותר - אי הוודאות של החיזויים עולה. שיטות מתורת המידע די נפוצות במחקר של רשתות עמוקות למשל:  
  
עקרון צוואר הבקבוק המידעי (של נפתלי תשבי): מציע ששכבות ברשת נוירונים שואפות לדחוס את המידע מהקלט תוך שמירה על המידע הרלוונטי לפלט. משמש לניתוח דינמיקת הלמידה והכללה של המודל.  
ניתוח מידע הדדי (Mutual Information): הערכת המידע ההדדי בין הקלט, השכבות הפנימיות והפלט מסייעת להבין כיצד מידע זורם ומעובד ברשת. זה הטכניקה שהמחברים משתמשים בה במאמר  
טכניקות רגולריזציה אינפורמטיביות: שיטות כמו information dropout שולטות בזרימת המידע במהלך האימון כדי לשפר חוסן והכללה של המודל.

אוקיי, אז המאמר מציע שיטה מבוססת מידע הדדי המשכללת MC dropout. במקום להשתמש ב dropout rate קבוע לכל השכבות (כלומר מה אחוז הנוירונים המחוסלים בשכבה) המחברים מציע לקבוע אותה (dropout rate) בתתלות במידת פגיעתה בזרימת המידע בשכבה. המטרה כאן היא לעשות את אובדן המידע בכל שכבה פחות או יותר קבוע. אם אובדן המידע הדדי גבוה(מקבוע אפסילון) מדי מקטינים dropout rate ואם זה נמוך מדי מגדילים אותו.

ד״א פגיעה בזרימת המידע בשכבה מחושבת דרך חישוב של המידע הדדי בין אקטיבציות של הקלט בשכבה לבין אלו של פלט השכבה. מתברר שזה די לא טריוויאלי והמאמר דן בהרחבה איך ניתן לעשות זאת.

https://arxiv.org/abs/2412.07169