⚡️🚀המאמר היומי של מייק 29.07.24: ⚡️🚀

Large Scale Dataset Distillation with Domain Shift

המאמר מציע שיטה מעניינת ודי מקורית לגנרוט דאטה מהתפלגות הנתונה על ידי דאטהסט מתויג. למשל בהינתן דאטהסט של תמונות D\_s המטרה היא ליצור דאטהסט (מתויג) גדול בעל התפלגות ה"מושרה" על ידי D\_s. המחברים טוענים כי השיטות הקיימות מתקשות לבנות(distill) דאטהסט גדול המשקף בצורה נאמנה את המאפיינים המהותיים של D\_s.

המחברים מציעים לגשת לבעיה זו עם גישה מעולם של domain adaption או DA בקצרה. בגדול מאוד DA היא תהליך של "התאמת מודל" במקרים בהם התפלגות הדאטה בזמן האינפרנס שונה מזו של הדאטה שעליה אומן המודל. התחום הזה עשיר בשיטות שחלקן די מורכבות מתמטיות ומערבות לרוב מינימיזציה של מרחק בין התפלגויות הדאטה (KL וכאלה).

למעשה המאמר המסוקר מתרגם את בעיית יצירת הדאטה לבעיית DA. התפלגות הדאטהסט שאנו מגנרטים ״ממנו״ D\_s משחק תפקיד של התפלגות המקור במקרה של DA (שעליו מאומן המודל ב-DA) ואילו התפלגות הדאטה המגונרט משחקת תפקיד של התפלגות היעד D\_t (כלומר זו של הדאטה שעליו מפעילים את המודל ב-DA). המטרה כאן לאמן מודל המקרב את ההתפלגויות האל.

אבל איך נחשב את ההתפלגויות האלו? המאמר מייצג את ההתפלגויות האלו על ידי התפלגות של האקטיבציות של השכבות השונות של הרשת. בפשטות עבור הדאטסט D\_s אנו מייצגים את התפלגות הדאטה על ידי וקטור הממוצעים ומטריצת קווריאנס של כל השכבות של המודל M\_s(מניחים שהם גאוסיים). בדיוק באותו האופן אנו מייצגים את ההתפלגות של הדאטה המגונרט.

אבל מה כאן M\_s ומה עושים כדי לקרב את התפלגות של הדאטה המגונרט להתפלגות הדאטה האמיתי? המודל M\_s אומן לשערך את ההתפלגות של הדאטהסט המתויג D\_s (המאמר לא מפרט איך M\_s מאומן בדיוק). למעשה האופטימיזציה מתבצעת על הדאטה המגונרט כאשר המודל M\_s נותר ללא שינוי. כלומר מתחילים מתמונות הנדגמות באקראי עם הלייבלים והמטרה היא לבצע מורד הגרדיאנט(gradient descent) על התמונות האלו במטרה לקרב אותם להתפלגות של D\_s.

עכשיו נשאלת השאלה מפונקציית הלוס כאן. כאמור בשלב הראשון אנו מאפטמים את התמונות המגונרטות במטרה למזער מרחק KL בין התפלגויות המשקלי המודל M\_s(נותר ללא שינוי) של D\_s (נותר קבוע לכל אורך הדרך) ולבין התפלגות של משקלי המודל M\_s עבור D\_t. המחברים מניחים ששתי התפלגויות אלו הם גאוסיים שעבורם מרחק KL ניתן לחישוב באופן מדויק בהינתם וקטורי תוחלות ומטריצות קווריאנס של D\_s ו- D\_t עם M\_s. איבר נוסף בלוס מנסה למקסם (=למזער עם סימן מינוס) הוא ההתפלגות המותנית של לייבל y בהינתן פיסת דאטה מג'ונרט (הרי אנו מגנרטים דאטה מתיוג). התיוג של כל פיסת דאטה מגונרטת נקבע מראש ולא משתנה במהלך האימון.

השלב השני הוא מזעור של מרחק KL בין ההתפלגות המותנית של הלייבלים של הדאטה המגונרט לבין זה של הדאטה מ-D\_s. בשביל כך מנצלים את הדאטה המגונרט מהשלב הראשון. מחשבים את התפלגות הלייבלים עבור הדאטה המגונרט הזה עם מודל M\_s ומאפטמים את הדאטה המגונרט במטרה לקרב את שתי ההתפלגויות האלו של הלייבלים (של הדאטה המגונרט ושל הדאטה מ-D\_s).

יש עוד לא מעט פרטים מעניינים על איך בדיוק מתבצע האימון (משתמשים בלא מעט מודלים לחישוב סטטיסטיקות המשקלים, עושים מיצוע מעריכים לסטטיסטיקות של הבאצ'ים וכדומה). המאמר לא כתוב מאוד ברור אבל הרעיון יפה.

https://dl.acm.org/doi/10.5555/3692070.3693400