המאמר היומי של מייק: 22.04.25  
Normalizing Flows are Capable Generative Models

בחרתי את המאמר הזה לסקירה כי הוא מכיל מכלול של גישות ושיטות שנדיר לפגוש היום במאמרי deep learning. הסיבה השנייה היא נוכחתה של שיטה גנרטיבית הנקראת Normalized Flow כלומר זרימות מנורמלות או NF בקצרה. זו שיטה שכמו GAN ו- VAE הפסידה בנוקאוט למודלי דיפוזיה בקרב על גישה גנרטיבית מובילה. עם זאת זו גישה מאוד מעניינת בעלת אפיון מתמטי מדויק ודי אינטואיטיבית. אז המחברים מנסים להחזיר את עטרה ליושנה ומציעים גישה מבוססת NF משולבת עם כמה כלים מתמטיים מעולם מודלי דיפוזיה ועוד טריק מתמטי נחמד הנקרא נוסחת Tweedie.

אז קודם כל מה זה NF? למעשה זו גישת אימון של מודלים גנרטיביים שמאמנת מודל של מיפוי 1-1 ערכי בין התפלגות פשוטה (כמו גאוסית סטנדרטית) לבין התפלגות הדאטה (נגיד דאטהסט של תמונות). מכיוון שהמיפוי הוא 1-1 ערכי אז הוא הפיך ועבור כל פיסת דאטה נתונה נוכל לחשב את נראותה (likelihood) ביחס למודל NF בצורה קלה (מפעילים את המיפוי ההופכי ומחשבים את הנראות של התוצאה לפי ההתפלגות הפשוטה). כמובן שהמימד של המרחב הלטנטי (המושרה על ידי התפלגות פשוטה) הוא בעל אותו המימד כמו מרחב הדאטה.

רוב מודלי NF עם מיפויים של הם הרכבה של מיפויים פשוטים ו1-1 ערכיים הפועלים על תת-קבוצה של מימדי הדאטה (שאר המימדים נותרים ללא שינוי). עבור תמונה למשל כל התפלגות אטומית כזו פעולת על כמה פיקסלים של התמונה. בד״כ מיפוי זה נבנה ממטריצות משולשות עליונות נלמדות (עם פונקציות לא לינאריות) מאחר ומטריצות אלו הן הפיכות וניתן להפוך אותן בצורה מאוד קלה (שזה מהווה יתרון עצום כאשר רוצים לחשב נראות של פיסת דאטה). בנוסף כל מטריצה כזו בנויה בצורה של חיבור residual כלומר היא מהווה סכום של הערכים הישנים והמיפויים שלהם.

אז המאמר בונה מודל NF בצורה כזו אך כאמור מציע כמה תופסות. התוספת הראשונה היא אימון המודל על דאטה מורעש. כלומר מאמנים את המודל על הדאטה שעבר הרעשה קלה (עם רעש גאוסי) שלטענת המחברים מעלה את הרובסטיות של המודל (די הגיוני בסך הכל). אבל כדי שהמודל לא ייצור לנו כתוצאה מכך פיסות דאטה מורעשות המחברים הציעו לנקות את הדאטה באמצעות נוסחת tweedie שלמעשה משערכת את התוחלת של דאטה מורעש (הדאטה שלנו) דרך דגימה של דאטה מורעש וגרדיאנט של לוג ההתפלגות של הדאטה המורעש ממושקל עם השונות. ככה בעצם מקבלים את הדגימה הנקייה מדגימה שהתקבלה אחרי אימון על דאטה מורעש.

והדבר האחרון שהמאמר עושה הוא שימוש ב Classifier Guidance שנמצא בשימוש כבוד במודלי דיפוזיה גנרטיביים. classifier guidance היא שיטה שמכוונת את תהליך הדגימה בעזרת מסווג חיצוני (classifier). במקום לדגום רק על סמך הרעש, המודל משלב את גרדיאנט ההסתברות של המסווג לתווית הרצויה, ובכך מעלה את הסיכוי שהדגימה הסופית תהיה שייכת למחלקה מסוימת. ניתן לעשות זאת ללא מסווג כאשר במקרה הזה אנו מזיזים את הדגימה בכיוון ההפוך המודל המגנרט דגימות לא מותנות (מודל דיפוזיה לא מותנה). אז המחברים מצאו דרך די אינטואיטיבית לדחוף את הרעיון הזה לתוך אימון מודל NF (בגדול מזיזים דגימה אחרי כל שלב ב-NF).

מאמר כיפי אך לא טריוויאלי בטח אם מנסים לצלול עמוק למתמטיקה אבל בתקווה העברתי את הרעיון הכללי…

https://arxiv.org/abs/2412.06329