המאמר היומי של מייק - 13.04.25  
ONE STEP DIFFUSION VIA SHORTCUT MODELS

המאמר מציע גישה מעניינת לאימון מודלי דיפוזיה גנרטיביים המהווה שכלול של שיטת flow matching (או FM בקצרה) שנהייתה הגישה המובילה לאימון מודלי דיפוזיה. למעשה המאמר מאמן מודל לשערך מסלול (בדרך כלל קו ישר שזה המסלול הכי פשוט אבל יש מאמרים שבוחרים צורות אחרות של המסלול) בין ההתפלגות הגאוסית (ההתפלגות הפשוטה) לבין התפלגות הדאטה (תמונות, וידאו או אודיו). המאמר טוען שבאמצעות השיטה המוצעת ניתן לגנרט דאטה באיטרציה אחת בלבד.

המודל מאומן לגנרט מהירות (גרדיאנט) של במסלול זה בכל נקודה t המסמנת כאן את עוצמת הרעש במסלול בין ההתפלגות הפשוטה (רעש טהור t = 0) להתפלגות של דאטה (t=1). אחרי שהמודל משערך מהירות זו ניתן לגנרט פיסת דאטה על ידי פתרון נומרי של משוואה דיפרנציאלית דרך הצבה של המהירות לשם. עבור מסלול לינארי המהירות הזו היא קבועה (נגזרת של קו ישר). לפעמים זה לא עובד כל כך טוב ומסלולים שנוצרים יוצאים לא לינאריים ודי מורכבים והדאטה שגונרט כתוצאה מכך לא מאוד איכותי.

המאמר מציע לבנות את המסלולים האלו לא בצורה לינארית אלא בצורה לינארית למקוטעין (סוג של ספליין לינארי) במקום להכריח את המודל ליצור מסלולים ממש לינאריים. התזוזה של נקודת דאטה בתת-מקטע תלויה רק בנקודה x\_t, t ובגרנוליריות הספליין d (ארחיב על זה אחר כך). תת-מסלולים אלו נקראים במאמר shortcuts והמודל מאומן לשערך אותם עם מה שנקרא consistency loss עליהם שכופה על המודל להיות ״עקבי״ בשני shortcuts עוקבים. לוס זה נגזר על ידי שילוב פשוט של הנוסחאות עבור ה-shortcuts העוקבים.

לאחר מכן המחברים משלבים את לוס ״העקביות״ זה עם הלוס הרגיל עבור FM (עם המסלול הישר). ניתן לבנות את המסלול מ-shortcuts בגרנולריות שונות של תת-קטעים לינאריים (כלומר עם מספר תת מקטעים שונים), אז האימון מנצל את זה ומאמן את המודל על גרנוריות שונות. כלומר בהינתן האיטרציה t(עוצמת הרעשה), דאטה מורעש וגרנולריות הספליין d המחברים מאמנים מודל המשערך את גודל הזזה של נקודת דאטה (shift) של תת מקטע הבא (כאמור יש d כאלו בסך הכל). לאחר מכן פותרים משוואה דיפרנציאלית כדי לקבל את ערך הדאטה בסוף התת-קטע. לאחר מכן שוב משערכים (באמצעות המודל המאומן) את הזזת נקודת דאטה שהתקבלה. ואז מפעילים את ה-consistency loss עבור שתי הזזות הדאטה.

מאמר די מעניין וכתוב בצורה נפלאה - מומלץ!

https://arxiv.org/abs/2410.12557