⚡️🚀המאמר היומי של מייק -17.10.24: ⚡️🚀  
EQUIVARIANT CONTRASTIVE LEARNING

היום נסקור מאמר שפורסם לפני שנתיים וחצי בנושא למידה ניגודית (contrastive learning). הנושא עצמו תמיד עניין אותי וסקרתי לא מעט מאמרים אבל חייב להגיד שבזמן האחרון שטף המאמרים על CL די נחלש. כאמור המאמר הזה שראה אור לפני שנתיים מציע שכלול לשיטה הקלאסית לבנייה של ייצוג דאטה (אמבדינג) באמצעות CL.

בגדול CL היא שיטה לבניית ייצוג של דאטה כאשר העיקרון המוביל הוא לקרב ייצוגי פיסות דאטה דומות(זוגות חיוביים) ולהרחיק ייצוגים של פיסות דאטה לא דומות (שליליים). זוגות דוגמאות חיוביים (במקרה של דאטה לא מתויג) נבחרות כאוגמנטציות שונות של דוגמא (עבור תמונות זה יכול להיות הזזה, סיבוב וכדומה) ואילו זוגות השליליים נבחרים באקראי מהדאטהסט.

אולם יש לא מעט בעיות עם הגישה הזו הקשורות לבחירת זוגות של דוגמאות חיוביות - למשל שני פאצ'ים באותה התמונה עלולים להכיל תוכן סמנטי שונה שלא נרצה לקרב את ייצוגיהם (הוצעו מספר פתרונות לסוגיה זו בעבר וחלקן סקרתי). בנוסף אולי היינו רוצים לקבל ייצוגים שונים (ולא מאוד קרובים) של טרנספורמציות מסוימות של אותה התמונה (נגיד סיבוב או הזזה) למשימת downstream ספציפית.

כלומר היינו רוצים להשרות יחס נתון T\_i בין ייצוגי התמונה ההתחלתית I ולייצוג התמונה אחרי טרנספורמציה T (נקרא לה I\_T). כלומר אנו רוצים לבנות ייצוג p כך ש:

p(T(I)) = I\_T(p(I))

וזה בדיוק מה שנקרא equivariance. למעשה CL הסטנדרטי הוא מקרה פרטי של equivariance שעבורן T\_i הינה טרנספורמצית זהות וזה נקרא אינווריאנטיות של הייצוג תחת טרנספורמציית T.

וזה בדיוק מה שהמאמר עושה. למעשה המחברים מציעים לאמן ייצוג ששומר על אינווריאנטיות עבור טרנספורמציות מסוימות (כמו בCL הסטנדרטי) ו אוכף בנוסף equivariance מוגדר לטרנספורמציות מקבוצה נתונה G המתאימה למשימת downstream שיש לנו ביד. כלומר לכל טרנספורמציה מ-G אנו מגדירים מראש את הטרנספורמציה ה-equivariant שלה (שיכולה להיות חברה ב-G גם כן) ומאמנים את הייצוג כך שהיחס ה-equivariance ביניהם יתקיים. מבחינה פרקטית הלוס הוא סכום משוקלל של הלוסים של CL הסטנדרטי ו ה-ECL.

מאמר חמוד - מחר או היום בערב אסקור את מאמר ההמשך שלו…

https://arxiv.org/abs/2111.00899