⚡️🚀המאמר היומי של מייק -08.11.24: ⚡️🚀  
Occam’s Razor for Self Supervised Learning: What is Sufficient to Learn Good Representations?

סקירה קצרה של מאמר המציע גישה חדשה ללמידה self-supervised או SSL בקצרה. אזכיר כי שיטת SSL מניחה שיש לנו דאטה לא מתויג ומתרטנו לאמן מודל מסוגל להפיק ייצוג חזק של דאטה. מה זה ייצוג חזק של דאטה, אתם שואלים? בד״כ הכוונה לכזה שניתן למנף אותו בצורה קלה (נגיד רק עם תוספת של שכבה לינארית) לבניית מסווג בעל ביצועים טובים.

כלומר כזה שיודע להפריד בין הקטגוריות השונות של דאטה בלי לדעת אותן בצורה מפורשת (למשל אנו יכולים לאמן מודל בצורת SSL על התמונות של ImageNet בלי להשתמש בתיוגים ואז לבדוק האם המודל הצליח ללמוד להפריד בין הקטגוריות השונות).

בד״כ SSL מבוצע עם שיטות של למידה ניגודית (contrastive learning) כאשר מטרתו מאוד בגדול היא לקרב ייצוגים של פיסות דאטה דומות (חיוביות) ולהרחיק את הייצוגים של פיסות דאטה לא דומות (שליליות). לרוב זוגות חיוביים נבחרים בתור אוגמנטציות שונות של אותה הדוגמא כאשר הזוגות השליליים הן דוגמאות שנבחרות באקראי. שיטות כאלו נחלו הצלחה די גדולה אבל דרשו דאטהסטים מאוד גדולים וגם משאבי אימון די משמעותיים (כי נדרש שם גודל באץ' די גדול כדי שהשיטה תעבוד טוב).

המאמר המסוקר מציע שיטה מאוד פשוטה ואינטואטיבית ל-SSL(תער אוקם). במקום לעבוד עם הייצוגים המאמר מאמן מודל לחזות את המספר של הדוגמא בדאטהסט. כלומר אם יש לנו 1000 דוגמאות מהדאטהסט יש לנו 1000 קטגוריות ומטרתנו לחזות קטגוריה של דוגמא מהייצוג הלטנטי שלה (המופק על ידי המודל המאומן). כלומר אחרי השכבה האחרונה של המודל מוסיפים שכבה עם מטריצה הממפה את הייצוג לקטגוריות (כלומר המספרים הסידוריים של הדוגמאות). ובסוף של לוס cross-entropy הסטנדרטי.

אז המאמר מוכיח שהשיטה עובדת לא רע לדאטהסטים יחסית לא גדולים (מעניין איך זה יעבוד לדאטהסט בגודל 10 מיליון). כמובן יש כמה טריקים באימון כמו soft labels אבל בגדול הרעיון די נחמד.

https://arxiv.org/pdf/2406.10743