המאמר היומי של מייק - 11.03.25  
Beyond Matryoshka: Revisiting Sparse Coding for Adaptive Representation

סקירה קצרה של מאמר המכליל שיטה להפקת ייצוג במימד נמוך של דאטה הנקראת Matryoshka embeddings. מה מיוחד בשיטה זו - היא מאפשרת לאמן את הייצוג הזה בכמה מימדים בו זמנים. כלומר במהלך האימון ייצוגים מכמה גדלים (נגיד 8ֿ, 16, 32, 64 ו-128) מאומנים באותו הזמן. השיטה מניחה דאטהסט מתויג של זוגות (x, y) כאשר x הוא פיסת דאטה ו-y הוא התיוג שלו.

ייצוגי מטריושקה מאמנים רשת עמוקה עם השכבה האחרונה (ראש) הממפה את הייצוג של דאטה לתיוג שלו. מה המיוחד במטריושקה הוא שהיא מאמנת בו-זמנית כמה וקטורי מיפוי (יחד עם המודל עצמו) למחרב התיוג כאשר כל מיפוי לוקח m\_i האיברים הראשונים מוקטור האמבדינג(השכבה האחרונה של המודל). בדוגמא שנתתי קודם מאמן בו זמנית וקטורי מיפוי בגדלים 8, 16, 32 ו-64. פונקצית הלוס הינה סכום של הלוסים עבור כל הוקטורים האלו - כלומר נוסף למודל עצמו אנו מאמנים 4 וקטורים בגדלים 8, 16, 32 ו- 64.

המאמר המסוקר מכליל את הגישה המעניינת הזו על ידי החלפתה בשני אלמנטים(של פונקצית לוס למעשבה). הראשון הוא sparse autoencoder או SE שבמקור מאומן למפות את ייצוג הדאטה, המופק על ידי המודל, למרחב בעל מימד מאוד גבוה אבל מאוד דליל ואז להחזיר אותו למרחב ייצוג המקורי. נציין כי המודל עצמו לא מאומן כאן אלא רק וקטורי המיפוי (של SE). האלמנט השני שמתווסף שהלוס ניגודי שבא להרחיק את ייצוגי הדאטה מקטגוריות שונות רחוק אחד מהם ולקרב את הייצוגים של פיסות הדאטה מאותה הקטגוריה.

ֿאז מה המטרה של SE כאן? להבדיל מהמטריושקה המקורית שמאמנת את האלמנטים הראשונים כאן אנו לוקחים top-k רכיבים של וקטור הייצוג אחרי האנקדור. הדקודר מאומן לשחזר את הוקטור המקורי רק עם top-k אלמנטים של הוקטור אחרי האנקודר. הבעיה הידוע עם SE היא הרכיבים של הוקטור אחרי האנקודר שלמעשה מתים - כלומר מקבלים ערכים קרובים מאוד לכל פיסות הדאטה.

כדי להתמודד עם בעיה זו החברים מציעים שני דברים. הדבר הראשון הוא הוספה לוסים עבור כמה ערכים של k ל- top-k של האנקודר לפונקציית לוס (במקור יש ערך k אחד). ככה אנו מאמנים אמבדינגס בכמה גדלים בדומה למטריושקה (חוץ מזה אין הרבה דמיון כי המטרה היא להפיק אמבדינג דליל). הדבר השני הוא הוספת של איבר המנסה לגרום לשגיאת השחזור עבור top-k של הרכיבים המתים (ערכים הכי נמוכים של וקטור הייצוג אחרי האנקודר) להיות קרוב לשגיאת השחזור של ה-top-k של הרכיבים הגדולים ביותר של ואותו הווקטור. אני לא הצלחתי לרדת לעומק דעתם למה זה עוזר.

בנוסף כאמור מוסיפים איבר של הלוס הניגודי לזה שמתואר בפסקה הקודמת….

טוב, נכון שהופיעה לנו המטריושקה בשם המאמר הדמיון בינו לבין המטרישקה המקורית די רופף. אבל המאמר די מעניין חוץ מזה….

https://arxiv.org/pdf/2503.01776