

לילה טוב חברים, היום אנחנו שוב בפינתנו deepnightlearners עם סקירה של מאמר בתחום הלמידה העמוקה  
היום בחרתי לסקירה את המאמר שנקרא:  
CoMatch: Semi-supervised Learning with Contrastive Graph Regularization  
שיצא ממש לפני כמה ימים

תחומי מאמר:

- (Semi-Supervised Learning (SmSL
- (Self-Supervised Learning (SSL

הוצג/יוצג בכנס: לא ידוע

כלים מתמטיים, מושגים וסימונים:

- (Self-Supervised Contrastive Learning (SSCL
- SSL/SmSL מבוסס על מינימיזציה של אנטרופיה
- SmSL המבוססים על גרף של דמיונות
- פסאודו לייבלים (תוצאת הרצת רשת סיווג על דאטה לא מתויג)
- יישור התפלגות (distribution alignment) עבור פסאודו לייבלים

תמצית מאמר: המאמר משלב 4 גישות פופולריות מעולמות של SSL ו-SmSL

1. רגולריזציה על בסיס עקביות: שיטה זו מבוססת על הנחה שההסתברויות של לייבל נתון, עבור דוגמא לפני ואחרי אוגמנטציה, אמורות להיות קרובות. למשל בדומיין של התמונות ההגיון מאחורי גישה זו הינו מאוד פשוט וטבעי: מכיוון שאוגמנטציה איננה משנה את התוכן של תמונה אלא רק את סגנונה, היא לא אמורה להשפיע על התפלגות פלט המסווג. עקרון זה ניתן לתרגם למשל למינימיזציה של קרוס-אנטרופי או מרחק ריבועי בין החיזויים של הדוגמה המקורית לגרסה שלה לאחר אוגמנטציה.

2. מינימיזציה של אנטרופיה של פלט המסווג: כאן אנחנו רוצים לבנות מסווג שמוציא "חיזויים בטוחים" לדוגמאות מהדאטה סט כלומר כאלו שלליבל אחד מקבל הסתברות גבוה משמעותית מכל האחרות. זה כמובן שקול למינימיזציה של אנטרופיה של פלט המסווג. בין השאר ניתן להשיג יעד זה ע"י מינימיזציה של פלט הרשת עבור דוגמאות לא מתויגות (בצורה מפורשת) או ע"י בנייה של פסאודו לייבלים בעלי אנטרופיה נמוכה על דוגמאות לא מתויגות ואימון של המסווג עליהן.

3. (Self-Supervised Contrastive Learning (SSCL: הנחת היסוד בגישה זו אומרת שייצוג חזק (במרחב במימד נמוך) של דאטה מסוגל להפריד בין זוגות של הדוגמאות דומות לבין זוגות של דוגמאות רנדומליות. אחת הצורות הפופולריות של פונקצית מטרה במאמרי SSCL נקראת InfoNCE. ניתן להראות כי ככל של InfoNCE קטן יותר, המידע הדדי בין הדוגמא במרחב המקורי לבין ייצוגה במרחב ממימד נמוך עולה. זה כמובן מצביע על אובדן פחות אינפורמציה בין הדאטה המקורי לבין ייצוגה כלומר הייצוג יהיה פחות לוסי ומייצג את הדאטה. חשוב לציין שהאימון מתבצע במרחב הייצוג ולא במרחב המקורי כלומר הלוסי מחושב על ייצוגים במרחב ממימד נמוך. לוס InfoNCE לוקח זוג של דוגמאות קרובות (למשל שתי אוגמנטציות של אותה דוגמא) ומספר דוגמאות רנדומליות ומנסה למקסם את היחס בין אקספוננט של דמיון של הזוג הקרוב לסכום הדמיונות בינו לבין דוגמאות רנדומליות.

4. SmSL המבוססים על גרף של דמיונות: כאן בונים גרף של דמיונות של דוגמאות מהדאטה סט כאשר קודקודים של דוגמאות קרובות (תחת איזושהי מטריקה-למשל במרחב המקורי או במרחב של לייבלים) מחוברים בקשת במשקל גבוה, כאשר הקודקודים של דוגמאות רחוקות מחוברים בקשתות בעלות משקל

נמוך או לא מחוברות כלל. אחר כך מנסים לאמן ייצוגים במרחב מימד נמוך של דוגמאות תוך כדי התחשבות ב"טופולוגיה של הגרף". במילים אחרות דוגמאות קרובות אחת לשנייה (מבחינת הגרף) יאומנו לקבל ייצוגים קרובים.

הסבר של רעיונות בסיסיים: המאמר מציע שיטה הנקראת CoMatch שלמעשה בנויה על שילוב של 4 גישות אלו. CoMatch מנצלת את הייצוג של דוגמאות במרחב לטנטי (מימד נמוך)  $Z$  ובמרחב הלייבלים  $Q$  ומבצעת אימון בהתבסס על שני גרפים של דמיונות הנבנים בהתבסס על קשרים בין דוגמאות במרחבים אלו. נציין כי  $Q$  הינו מרחב הפלטים של רשת הסיווג כלומר הוא מכיל וקטורי הסתברויות של הלייבלים. אז איך זה בעצם נעשה? קודם כל בואו נבין את המבנה של פונקציה הלוס של CoMatch.

פונקציה לוס: נתחיל מזה שנזכר CoMatch הינה שיטה של SmSL כלומר יש לנו דאטה סט עם דוגמאות מתויגות הנקרא  $X$ , ודאטה סט של דוגמאות לא מתויגות  $U$ . נסמן את הגרף שנבנה מעל המרחב הלטנטי  $Z$  ב  $G\_emb$ , והגרף על מרחב לייבלים  $Q$  יסומן ב  $G\_lab$ . עכשיו נוכל לעבור לתיאור הרעיונות העיקריים של המאמר: המאמר מציע לאמן 3 רשתות:

- רשת המקודדת  $f$  שבונה ייצוג מקדים של הדאטה המשמשת גם כהשלב הראשון לבניית של הייצוג  $z$  גם לסיווג עצמו
- רשת שבונה ייצוג במרחב הלטנטי  $Z$  שמופעלת אחרי  $f$  המסומנת ע"י  $g$  (הייצוגים תמיד מנורמלים)
- רשת סיווג  $h$  שמטרתה להוציא וקטור הסתברויות של לייבלים (גם מופעלת אחרי  $f$ )

עכשיו נדבר על פונקציית הלוס של המאמר המורכבת מ 3 חלקים:

1.  $L\_x$ : קרוס-אנטרופי לוס רגיל על דוגמאות מתויגות. כאן הלוס מחושב על דוגמאות מתויגות שעברו אוגמנטציה חלשה כלומר בין החיזוי של הדוגמא אחרי אוגמנטציה לבין הלייבל של התמונה המקורית)
2.  $L\_ucls$ : קרוס אנטרופי לוס על פסאודו לייבלים (הרכים!!) של דוגמא לא מתויגות לבין החיזוי של הרשת עבור אותה דוגמא שעברה אוגמנטציה חזקה. כאן רק פסאודו לייבלים מעל סף מסוים נלקחים בחשבון בחישוב הלוס בשביל לא לקנוס את המודל על הדוגמאות שלא הצלחנו לבנות להם פסאודו לייבל "אמין", כלומר עם אנטרופיה נמוכה ([FixMatch](#)). על איך בונים את הפסאודו לייבלים האלו נדון בפרק הבא
3.  $L\_uclr$ : הלוס הקונטרסטי (בסגנון InfoNCE) הבנוי על גרפי הדמיונות על מרחבי  $Z$  ו-  $Q$ . נסביר את המבנה של לוס זה בהמשך.

כעת בואו נתעמק באיך בונים את הפסאודו לייבלים  $q\_b$  הנחוצים לחישוב של  $L\_ucls$ . יצירה של פסאודו לייבלים: קודם כל עבור דוגמאות מתויגות הפסאודו לייבל שווה ללייבל האמיתי שלהן. עבור דוגמאות לא מתויגות קודם מפעילים אוגמנטציה חלשה ומחשבים את ההתפלגות החזויה של הלייבלים שלהן. לאחר מכן מבצעים יישור התפלגות (DA) המיועד בשביל לא לתת להתפלגות הלייבלים לקרוס לתת-קבוצה של הלייבלים. כאן מחשבים ממוצע נע  $p\_av$  (המהווה בעצם שיעור של שכיחות הלייבלים בדאטה סט) על כל החיזויים של הדוגמאות הלא מתויגות ומחלקים את וקטור ההסתברויות החזויות עבור דוגמא נתונה  $p\_w$  ב-  $p\_av$  נציין שלהדביל מ- [ReMixMatch](#) וקטור שכיחויות הלייבלים הנגזר מהדוגמאות המתויגות לא נלקח בחשבון כאן. אחר כך מחשבים גם את הייצוגים הלטנטיים  $z\_w$  של הדוגמאות (אחרי אוגמנטציה חלשה) ע"י העברתם דרך הרשת המקודדת  $f$  ורשת הייצוג  $g$ . שומרים  $z\_w$  ביחד עם  $p\_w$  במאגר של דוגמאות  $B$ . בשביל למצוא את הפסאודו לייבל  $q\_b$  של דוגמה לא מתויגת, המאמר מנסח את בעיית אופטימיזציה עבור  $q\_b$  עם פונקציה מטרה המורכבת מסכום (עם מקדמים המסתכמים ל 1) של שני מחוברים (נזכיר ש  $q\_b$  זה בעצם התפלגות על מרחב הלייבלים):

- סכום הריבועים של המרחקים של פלט הרשת המסווגת  $h$  עבור דוגמא  $u\_w$  לבין הפלטים של כל הדוגמאות  $u\_wk$  מ-  $B$ , כאשר כל מרחק ממושקל בדמיון המנורמל  $a\_k$  בין הייצוג הלטנטי  $z\_w$  של  $u$  לבין הייצוג הלטנטי של  $u\_wk$ . המטרה של האיבר הזה לקרב את התפלגויות של הלייבלים

עבור דוגמאות קרובות. כאן דמיון בין הייצוגים מוגדר כאקספוננט של המכפלה הפנימית בין הייצוגים (המנורמל בסכום של כל הדמיוניות עבור הדוגמאות מ B) המרחק הריבועי בין החיזוי של הדוגמה  $p_w$  לבין  $q_b$

לבעיית אופטימיזציה זו יש פתרון מדויק וזה למעשה סכום ממושקל עם  $p_w$  ו  $p_{wk}$  כאשר המקדם של לפני  $p_{wk}$  זה הדמיון  $a_k$

החלק האחרון בפאזל שטרם התייחסנו אליו הינו הלוס  $L_{uctr}$  המבוסס על גרפי דמיון מעל מרחבי ייצוגים Z והלייבלים Q.

מבנה של  $L_{uctr}$ : לכל באטץ' בונים גרף  $G_{lab}$  מעל מרחב Q כאשר משקל הקשת בין דוגמאות (קודקודים) מוגדרת ע"י הדמיון (מכפלה פנימית) בין הפסאדו לייבלים של הדוגמאות (משקל של קשת עצמית מוגדרת להיות 1). אם דמיון זה מתחת לסף הקודקודים של דוגמאות אלו לא מחוברים. לאחר מכן בונים גרף  $G_{emb}$  מעל מרחב הייצוגים כאשר הקשת העצמית של קודקוד מוגדרת כדמיון בין הייצוג של שתי אוגמנטציות חזקות של דוגמא u (הדמיון כאן זה בעצם מכפלה פנימית בין הייצוגים). הקשת בין כל זוג אחר של קודקודים מוגדרת כדמיון בין הייצוג של הדוגמא (לאחר אוגמנטציה חזקה) לבין הייצוג של הקודקוד השני. לאחר מכן מנרמלים את הקשתות עבור שני הגרפים. ועכשיו בא הקטע המגניב של המאמר הזה (לפחות בעיניי) וזה בניית הלוס המשלב את שני הגרפים האלו. לוס זה מורכב משני חלקים:

- הלוס הקונטרסטי בין הקשתות העצמיות של  $G_{lab}$ . בדומה ללוסיים קונטרסטיים דומים הוא דוחף את המודל לתת ייצוגים דומים לאוגמנטציות שונות של אותה דוגמא.
- דוחף את המודל לבנות ייצוגים דומים עבור לדוגמאות דומות עם פסאדו לייבלים דומים (!!)(כלומר קשת בעלת משקל גבוה ב-  $G_{lab}$ ). לדעתי זו אחת הנקודות הכי חשובות במאמר ולדעתי הסיבה לכך ש CoMatch הצליחה להגיע לביצועים טובים.

הישיג מאמר: מראים שיפור בביצועים בכמה משימות SmSL קלאסיות מעל שיטות עכשוויות כמו [FixMatch](#) ו- [MixMatch](#)

דאטה סטים: CIFAR100, STL10

לינק למאמר: <https://arxiv.org/pdf/2011.11183v1.pdf>

לינק לקוד: לא מצאתי

נ.ב. המאמר מציע שילוב אלגנטי של 4 שיטות אימון המקובלות היום ב SmSL. מאוד אהבתי את השילוב של גרפי דמיון מעל מרחבי ייצוג ומרחב הלייבלים בחישוב של הלוס הקונטרסטי. עם זאת המאמר הראה את עליונות של CoMatch רק על שני דאטה סטים יחסית קלים. הייתי רוצה לראות את ביצועיה של גישה זאת לדאטה סטים יותר מורכבים ומקוואה שזה יבוא בהמשך.

#deepnightlearners