

דרישות להגשת פרוייקט גמר

שיטות מתקדמות בעיבוד תמונה רפואי 80303 / 6.2021

1. הפרוייקט יתבצע ע"י שני סטודנטים האחראים במשותף על ההגשה עד לתאריך 23.7.21.
2. בהרצאה האחרונה, 30.6.21, כל הצוותים יציגו את הפרוייקט שלהם (חובה!) בכיתה. משך ההצגה יהיה 6-8 דק'. ההצגה הינה חלק מציון הפרוייקט.
3. הקפידו להגיש את הקוד (ניתן להעלות ל- Github) כ- **Jupyter notebook** (אפשר ב colab) כאשר כל בלוק של קוד יכיל הסברים ברורים.
4. כל צוות יירשם בקובץ משותף בו יציין את שמות הסטודנטים, הבעייה ומאגר הנתונים שנבחרו. המטרה היא למנוע ככל הניתן כפילות במקורות הנתונים.

הצגת הפרוייקט

המצגת צריכה להיות לכל היותר 10 שקפים ועליה לכלול:

1. כותרת, מחברים
2. מבוא: סעיף זה מציג את הפרוייקט. מדוע הוא חשוב או מעניין. מה הבעיה שפותרים, תיאור ה-input/output וכיצד נרכש.
3. תיאור מסד הנתונים – תמונות:
 - a. כמה דוגמאות ישנם וכיצד אתם מתכוונים לחלק אותם בין שלבי האימון. האם יש מספיק תמונות? נדרש לכלול הפניה למקור מערך הנתונים. כמו כן, יש להעלות את המסד לענן כלשהוא (למשל גיט)
 - b. האם נדרש עיבוד מקדים? מה לגבי נורמליזציה או הגדלת נתונים (אוגמנטציה)? מה הרזולוציה של התמונות שלך? הציגו כמה דוגמאות והסבירו אותם.
 - c. יש להסביר את ה meta-data של התמונות וכיצד יעשה בהם שימוש.
4. תיאור גישת הפתרון:
 - a. תארו את הצעדים הנוכחיים שביצעתם, האלגוריתמים שאתם מתכוונים ליישם. אם התחלתם יישום ראשוני של המודל ו/או חקירת הנתונים, תארו אותם.
 - b. תארו במדויק את העבודה שנותרה - תיאור מודל ואסטרטגיית האימון (loss func. למשל), קריטריונים לביצועים וכו'. מובן שהמודל עשוי להשתנות בהמשך.

הנחיות לדו"ח המסכם

1. מבוא: בהקדמה עליכם לתת סקירה לנושא הפרוייקט – הרחבה של סעיפים 1,2 מהצגת הפרוייקט. הסבר מפורט יותר של הדאטה, אלו פעולות בצעתם כדי להשתמש בו לאימון.
2. עבודה קשורה related work : כאן תדונו בקצרה ב 2-3 עבודות הרלוונטיות לפרוייקט מהספרות.

3. שיטות: בחלק זה תפרטו את שיטות הלמידה העמוקה שנוסו במהלך הפרויקט. יש ליישם ולהעריך את הדברים הבאים:

- a. היכן שרלוונטי - (בעיקר לסיווג, אבל לא רק) תוצאת base-line של למידה לא עמוקה (יכול להיות למשל גישת למידת מכונה קלאסית כמו SVM לסיווג).
- b. גישת למידה עמוקה מתקדמת (כלומר הגישה שלכם לפרויקט הגמר) - עליכם להיות מפורטים לגבי הגישה, ולתאר את המודל שבחרתם: ארכיטקטורה, כלומר כמה שכבות וסוגם? איזה קצב למידה בחרת? איזה שיטת אופטימיזציה? איך נראית הארכיטקטורה שלכם? פונקציית המחיר? איך עשיתם כוונן של היפר-פרמטרים? **דמיינו שמי שקורא את הדו"ח ינסה לשחזר במדויק את התוצאות שלכם מבלי להתבונן כלל בקוד שלכם - האם יהיה מסוגל לעשות זאת?**

4. ניתוח/אנליזה

- a. ניתוח השיטות שתוארו בסעיף הקודם. השוו את השיטות.
- b. איזה מודל עובד הכי טוב? למה? אם יישמתם מודלים נוספים, אך גיליתם שהן לא הצליחו, הקפידו לכלול גם את הניתוחים שלהם כאן. הניתוח חייב להכיל לפחות את הנקודות הבאות: מדדי טיב/ביצועים רלוונטיים, בהתאם לבעיה. בבעיית סגמנטציה ניתן להשתמש במדד Dice להערכת איכות. בבעיית סיווג יש לספק את אחוזי הניבוי, מטריצת ה confusion, עקומת ROC.
- c. ויזואליזציות: יש לספק בדוח את עקומות הלימוד/הכללה וכו'. דוגמאות למוצא המודל.
- d. דיון מדוע מודל אחד פועל טוב יותר מאחר. השתמשו בעקומות לימוד וכל ראיות אחרות מתוצאות הניסוי כדי להסביר את ההבדל.

שימו לב שגישת הפתרון שלכם בפרויקט לא צריכה לתת ביצועים טובים יותר מהספרות וכו'. המטרה העיקרית היא היכולת שלכם להסביר מה עשיתם, תהליך בחירת המודל שלכם, הערכת הביצועים, טיוב המודל וכו'.

5. דיון

בהתבסס על התוצאות והעבודה שביצעתם, איזו עבודה נוספת נדרשת בפרויקט? אילו גישות אחרות הייתם מנסים בהינתן יותר זמן? מה המסקנות שתוכלו להסיק מהתוצאות שלכם?

6. הפניות references : ביבליוגרפיה

דוגמה לדו"ח (להתייחסות): http://cs230.stanford.edu/projects_fall_2020/reports/55817960.pdf