

05/26/2025

פרויקט מסכם בקורס

שיטות למידת מכונה לניתוח אותות fMRI

צוות הקורס:

מרצה: ד"ר ארז סימוני

כללי הגשה:

- את התרגיל יש לעשות אך ורק בזוגות. כל חריגה , חייבת את אישור הצוות.
התרגיל יוגש לתיבה שתפתח באתר , ויוגש רק מאחד מבני הזוג.
יש להעלות קובץ יחיד zip, אשר מכיל שני קבצים:
(1) קובץ מסכם , עם הסברים מפורטים, יחד עם הגרפים של התוצאות, ודיונים בתוצאות.
(2) יש לצרף את הקוד, כאשר הוא מתועד ונוח לקריאה, בקובץ נפרד.

מטרת התרגיל:

זיהוי סרטים מתוך סריקות fMRI 7T באמצעות מסווגים לינארים

מהיכן מגיעים הנתונים לפרויקט, ומה הניסוי?

170 נבדקים מה – Human Connectome Project (HCP) אשר צפו ב-14 סרטים (קליפים) קצרים באורך של 50 שניות, כאשר ביניהם יש מרווח של rest באורך 20 שניות.

בתרגיל אנחנו נעבוד על דאטה מהמוח של 170 נבדקים, מ-3 רשתות שצפו ב-14 סרטים (3 קבצים - כל רשת קובץ אחד), ודאטה של "מנוחה" בין הסרטים מכל איזור של 20 שניות (3 קבצים).

שלושת רשתות (vis), רשת הקשב (DAN), ומרשת ברירת המחדל (DMN).

קישור לקבצים הדאטה להורדה:

https://drive.google.com/drive/folders/195_f-f0sq-07-UHHiFZId7mkEI2OuM4-?usp=drive_link

מבנה הנתונים:

170 14 50 34

מבנה קובץ/משתנה הקשור לסרטים:

ניקח לדוגמא את הקובץ data_movie_vis.mat (רשת ויזואלית)

לדוגמא טעינתו של קובץ זה – מתקבלת מטריצה בעל הגודל $170 \times 14 \times 50 \times 34$

170 נבדקים

14 סרטים (קליפים)

50 : 50 שניות מסוף הסרט (הדגימה האחרונה (50) היא למעשה סוף הסרט)

34 איזורים

מבנה קובץ/משתנה הקשור לזמני המנוחה בין הסרטים:

ניקח לדוגמא את הקובץ data_movie_vis.mat - מתקבלת מטריצה בגודל: $170 \times 14 \times 19 \times 34$

170 נבדקים

14 קטעי מנוחה (בין קליפים)

19 : 19 שניות בין סרט לסרט

34 איזורים

מהלך הפרויקט:

1. חלק ראשון (סיווג 14 סרטים לפי 5 שניות האחרונות והראשונות של כל סרט)

בחלק זה אתם אמורים לסווג איזה סרט ראה כל נבדק, לפי ה-5 שניות האחרונות והראשונות של כל סרט).

לדוגמא : אם לרשת ה-VIS יש 34 איזורים, עבור נבדק מספר אחד וקליפ מספר אחד, אתם מקבלים מטריצה שהיא 50×34 . מכאן אתם מחלצים את ה-5 שניות האחרונות, ואת ה-5 שניות הראשונות ומקבלים מטריצה של 5×361 עבור סוף הסרט, ו- 5×361 עבור תחילת הסרט (זה לא בדיוק תחילת הסרט – אלא 50 שניות מסיום הסרט).

כל מטריצה כזאת הופכים לוקטור ("משטחים" את המטריצה), ומצמידים לה לייבל של הסרט, כך בונים מטריצה X שכל שורה שלה היא המטריצה המשוטחת, והאיבר המתאים בוקטור Y הוא הלייבל של הסרט.

- (1) יש לאמן את המסווגים מסווגים : Naïve Bayes, QDA, LDA, SVM, KNN לזיהוי בין 14 הסרטים לכל איזור ואיזור, מהסוף ומהתחלה. (נסו לאמן עם leave-one-out-cross-validation ואם זה איטי עברו ל k-fold-cross-validation)
- (2) יש להציג את תוצאות accuracy ב bar plot עבור כל ה 3 האיזורים (מהתחלה ומהסוף), עבור כל מסווג. לכל איזור (עמודה) יהיה "2 מקלות" (המציינים סיווג של תחילת הסרט וסוף הסרט).
- (3) יש להציג את טבלת ה confusion matrix – לכל 6 המצבים – רק עבור המסווג הטוב ביותר.

2. חלק שני (סיווג 14 מקטעי המנוחה לפי 5 שניות האחרונות והראשונות של כל מקטע מנוחה)

בחלק זה אתם אמורים לסווג איזה סרט ראה כל נבדק, לפי ה-5 שניות האחרונות והראשונות של מקטע המנוחה (19 שניות) שמגיע לאחר כל סרט.

מתוך המידע ניתן לחלץ 14 מקטעי מנוחה (כל מקטע 19 שניות) מ-170 נבדקים מתוך 3 הרשתות. בחלק זה אתם נדרשים לבדוק ביצועים של סיווג המקטעים האלה.

יש לבצע את הסיווג כל פעם על מקטע של 5 שניות ראשונות ואחרונות מתוך ה-19 שניות בקטע המנוחה באותו אופן כפי שמתואר בחלק הראשון.

- (1) יש לאמן את המסווגים מסוג : Naïve Bayes, LDA, QDA, SVM, KNN לזיהוי בין חמשת מקטעי המנוחה מ-5 האיזורים, באמצעות וקטור הווקבלים מאמצע המנוחה. יש להציג את תוצאות accuracy ב bar plot עבור כל ה 5 האיזורים. לכל איזור (עמודה) יהיה 5 "מקלות" (המציינים

סיווג).

השתמשו leave-one-out-cross validation

(2) עבור 3 האיזורים – יש להציג את טבלאת ה confusion matrix – רק עבור המסווג הטוב ביותר.

בנוס:

חשבו על features אחרים לסיווג – או מסווגים אחרים - ובדקו מדגמית האם תוכלו לקבל סיווג טוב יותר
? גבו את תוצאתכם בגרפים מתאימים

חלק שלישי סיכום הפרוייקט

דונו בתוצאות – בהתייחס לשלושה איזורים שונים של המוח , והסבירו את התוצאות של תחילת הסרט לעומת סוף הסרט , וכן תחילת המנוחה וסוף המנוחה.