

פרויקט גמר בפייתון - זיהוי בלבול מאותות EEG

הדס שיף 209389881
יותם נצר 322217340

מבוא

במסגרת הפרויקט ניתחנו גלי מוח (EEG) אשר נאספו מאוכלוסיית סטודנטים במהלך צפייה בסרטונים חינוכיים. הראו לתלמידים 10 סרטונים שונים, 5 הנחשבים "קלים" על נושאים בסיסיים, ו 5 הנחשבים "קשים". מהמחקר התקבלו נתוני ה-EEG של 10 תלמידים במהלך הלמידה והמשוב שלהם לגבי רמת הבלבול מהנושא אותו למדו. נתוני ה-EEG נדגמו כל 0.5 שניות מהאונה הקידמית. (סה"כ ישנם 15 שדות שונים הכוללים את האותות ואת הזיהוי של הסרטונים)

מטרת המחקר הייתה לבדוק האם ישנם הבדלים מובהקים בפעילות המוחית במצבים שבהם הסטודנטים צפו בסרטונים המוגדרים כ"מבלבלים" לעומת מצבים שבהם הסרטונים לא הוגדרו ככאלה.

בנוסף, הפרויקט חקר האם קיימים הבדלים משמעותיים בפעילות המוחית כאשר הסרטונים סווגו כמבלבלים על ידי מערכת חיצונית (כלומר, על ידי מבצעי המחקר), לעומת מקרים שבהם הסטודנטים עצמם דיווחו באופן סובייקטיבי על תחושת בלבול.

המטרה העיקרית של הפרויקט הייתה לפתח כלי שיאפשר זיהוי תחושת בלבול באופן אובייקטיבי, תוך שימוש בניתוח אותות EEG, וכך לספק תובנות על דרכים לשיפור חוויות למידה מותאמות אישית.

שיטות עבודה

איסוף נתונים

הנתונים כללו:

- אותות EEG גולמיים: נמדדו באמצעות אלקטרודה אחת לאורך הצפייה בסרטונים.
- תיוגים חיצוניים: כל סרטון סווג כ"מבלבל" או "לא מבלבל" על פי החלטת החוקרים.
- תיוגים סובייקטיביים: הסטודנטים התבקשו לדווח על תחושת בלבול לאחר הצפייה בכל סרטון.

עיבוד נתונים

- ניקוי הנתונים: בוצע סינון ראשוני להסרת ערכים חסרים ורשומות כפולות. בוצע יישור מדויק של הרשומות להבטחת עקביות בין תוויות הבלבול לאותות EEG המתאימים.
- שילוב תיוגים: אותות EEG חולקו לשתי קבוצות עיקריות:
 - "מבלבל" לפי סיווג חיצוני.
 - "מבלבל" לפי דיווחי הסטודנטים.

ניתוח נתונים

- מבחני T למדגמים תלויים: בוצעו השוואות סטטיסטיות בין האותות משתי הקטגוריות. המבחנים נועדו לבדוק אם ישנם הבדלים מובהקים בין "מבלבל" ל"לא מבלבל", בין אם מדובר על תיוג חיצוני או סובייקטיבי.
- ויזואליזציה: נוצרו גרפים והיסטוגרמות שהציגו את ממוצעי האותות, פיזור הנתונים וההבדלים בין הקבוצות.
- למידת מכונה: פותח מודל עץ החלטה (Decision Tree) כדי לזהות דפוסים באותות המוח ולהבדיל בין מצבים של בלבול וחוסר בלבול. המודל הוערך על בסיס מדדי דיוק, רגישות ושלמות.

תוצאות ודיון

- בדיקה ראשונית של הנתונים אישרה כי אין ערכים חסרים או רשומות כפולות.

- הנתונים הכילו 100 אירועים של סטודנטים הצופים בסרטונים.
 - ניתוח על פי תוויות חיצוניות:
 - מבין כל עמודות הנתונים, רק גלי המוח של Theta עברו את בדיקת הנורמליות.
 - תוצאות מבחן הז הצביעו כי אין הבדלים סטטיסטיים מובהקים בין הגלים בפעילות בזמן צפייה בסרטונים "מבלבלים" לעומת "לא מבלבלים".
- פרמטרים סטטיסטיים:
 t -סטטיסטי: 0.1081
 P-value: 0.9143
- תוצאות מודל עץ ההחלטה:
 - דיוק ממוצע: 0.5286
 - דיוק (Precision) ממוצע: 0.5364
 - רגישות (Recall) ממוצעת: 0.5342

- ניתוח על פי תוויות סובייקטיביות של הסטודנט:
 - מבין כל עמודות הנתונים, רק גלי המוח של Theta ושל Alpha1 עברו את בדיקת הנורמליות.
 - תוצאות מבחן הז הצביעו על הבדלים סטטיסטיים מובהקים בין הגלים בפעילות בזמן צפייה בסרטונים "מבלבלים" לעומת "לא מבלבלים".
- פרמטרים סטטיסטיים עבור Theta:
 t -סטטיסטי: 8.5768
 P-value: 0.0000
- פרמטרים סטטיסטיים עבור Alpha1:
 t -סטטיסטי: 5.8383
 P-value: 0.0000
- תוצאות מודל עץ ההחלטה:
 - דיוק ממוצע: 0.7355
 - דיוק (Precision) ממוצע: 0.7448
 - רגישות (Recall) ממוצעת: 0.7421

תובנות מרכזיות

1. תוויות בלבול מוגדרות מראש גילו יכולת מוגבלת בזיהוי הבדלים בפעילות עצבית.
2. תוויות מוגדרות על-ידי המשתמש הראו הבדלים משמעותיים בפעילות Theta ו-Alpha1.
3. מודל עץ ההחלטה הציג שיפור דרמטי בביצועים עם תוויות משתמש.

כל אלו מרמזים על כך שתחושת הבלבול האישית משפיעה באופן חזק יותר על פעילות המוח מאשר סיווג חיצוני של תוכן הסרטון, ולכן יש מקום לפתח מודלים מותאמים אישית שמתחשבים במדדים סובייקטיביים.

השימוש באותות EEG ככלי לניטור תחושת בלבול מציג פוטנציאל משמעותי בשיפור חוויית הלמידה, במיוחד בסביבות חינוכיות מותאמות אישית.

נדרש מחקר נוסף כדי להבין טוב יותר את דפוסי ה-EEG במצבים של בלבול ולשפר את הדיוק של מודלי החיזוי.