**דו"ח – פרויקט מסכם תכנות פייתון**

**מגישות: אריאל צומן, 211778410 | מאי מועלם, 207105453 | לינוי אלבז, 322212051**

מבוא ורקע:

מחלת פרקינסון היא מחלה נוירולוגית פרוגרסיבית הפוגעת במערכת העצבים המרכזית. היא מתאפיינת בירידה ברמות הדופמין במוח. אחד מהתסמינים הנלווים למחלה הוא ירידה ביציבות הקול, המתבטאת בשינויים בקצב ובקוליות הדיבור. תופעה זו יכולה להוביל לדיבור לא ברור, לתנועות לא רצוניות של הגרון או של שפת הגוף, ולעיתים אף לבעיות נשימה. בנוסף, חולי פרקינסון נוטים לחוות ירידה בעוצמת הקול, מה שמקשה עליהם לתקשר בצורה אפקטיבית. תופעה זו נקראת לעיתים "אפקט הרעש הפנימי", שבו החולה מדבר בקול חלש מבלי להיות מודע לכך.

**שאלת המחקר:** האם קיימים הבדלים משמעותיים ביציבות ובעוצמת הקול בקרב חולי פרקינסון בהשוואה לאוכלוסייה בריאה?

**השערה:** חולי פרקינסון יציגו יציבות מופחתת בטון ועוצמת קול נמוכה יותר בהשוואה לאוכלוסייה בריאה.

סקירת הData:

הData כללה מספר משתנים, כאשר כל המשתנים מייצגים מדדים שונים שקשורים לתכונות הקול. העמודות הרלוונטיות לנו היו העמודות שקשורות ליציבות ועוצמת הקול:

* MDVP:Fo(Hz) - תדר היסוד הוא התדר הבסיסי ביותר של הקול, והוא נקבע על פי קצב הרטט של מיתרי הקול.
* MDVP:Fhi(Hz) - מייצג את התדר המקסימלי של הקול, ומשקף את הנקודות שבהן מיתרי הקול רוטטים בקצב המהיר ביותר במהלך הדיבור.
* MDVP:Flo(Hz) - מייצג את התדר המינימלי של הקול, ומשקף את הנקודות שבהן מיתרי הקול רוטטים בקצב האיטי ביותר במהלך הדיבור.

הנתונים מצוינים ביחידות של הרץ, (Hz) כאשר 1 הרץ מתייחס לתנודה אחת בשנייה. התדר מודד את מספר התנודות של גל, ובמקרה זה- הקול, בכל שנייה. ככל שהתדר גבוה יותר, כך הצליל נשמע גבוה יותר , ולהפך.

שיטות:

תחילה, הסרנו את העמודות שאינן רלוונטיות לשאלת המחקר. לאחר מכן, התחלנו בתהליך ניקוי הנתונים, שכלל הסרת שורות ועמודות עם ערכים חסרים. המשכנו בנרמול הערכים בעמודות לטווח של 0-1, במטרה לצמצם את השונות בין הערכים ולהביאם לסקאלה אחידה. בנוסף, הסרנו נתונים חריגים והמרנו עמודות עם טיפוס טקסטואלי לערכים מספריים.

לאחר ניקוי הנתונים, ביצענו אנליזות על העמודות השונות, כולל חישוב סטטיסטיקות תיאוריות כגון: מספר השורות בעמודה, ממוצע, ערכים מינימליים ומקסימליים, אחוזונים וסטיית תקן. בדקנו גם אם הערכים בכל עמודה מתפלגים באופן נורמלי.

ערכנו השוואת ממוצעים בין שתי הקבוצות: בריאים מול חולי פרקינסון, עבור העמודות הרלוונטיות, באמצעות מבחן t דו-מדגמי. כמו כן, יצרנו מטריצת מתאם בין כל זוג עמודות כדי לבדוק את הקשרים בין המשתנים. לבסוף, ביצענו רגרסיה לוגיסטית שלמעשה בנתה מודל חיזוי פשוט כדי לנבא משתנה בינארי.

תמונה שמכילה צילום מסך, טקסט, מלבן, ריבוע

התיאור נוצר באופן אוטומטיתובנות מרכזיות:

הפונקציה מחשבת את מטריצת המתאם עבור העמודות: MDVP:Fo , MDVP:Fhi ו,MDVP:Flo באמצעות מפת חום. כל תא במפה מייצג את המתאם בין שני משתנים, כאשר הערכים נעים בין 1 ל1-:

* **1:** קשר חזק וחיובי (ככל שמשתנה אחד עולה, השני גם עולה).
* **-1:**  קשר חזק ושלילי (ככל שמשתנה אחד עולה, השני יורד).
* **0:** אין קשר.

ניתן לראות כי המתאם, עוצמת הקשר והכיוון בין שני משתנים, בין MDVP:Floל MDVP:Foהוא הקשר החזק והחיובי ביותר (=0.6). בניגוד למתאם בין MDVP:Flo ל MDVP:Fhiשהוא החלש ביותר, אך עדיין קשר חיובי (=0.08).

תמונה שמכילה צילום מסך, טקסט, מלבן, תרשים

התיאור נוצר באופן אוטומטיגרף 1 MDVP:Flo(Hz) (התדר המינימלי של הקוד):

ממוצע הקבוצה הבריאה הוא 0.46

ממוצע הקבוצה של חולי פרקינסון הוא 0.24

הממוצע של עבור הבריאים גבוה משמעותית מזה של חולי פרקינסון.תמונה שמכילה צילום מסך, טקסט, מלבן, תרשים

התיאור נוצר באופן אוטומטי

גרף 2 MDVP:Fhi(Hz) (התדר המקסימלי של הקוד):

ממוצע הקבוצה הבריאה הוא 0.25

ממוצע הקבוצה של חולי פרקינסון הוא 0.18

גם כאן ניכר הבדל בין שתי הקבוצות, הבריאים מציגים ממוצע גבוה יותר.תמונה שמכילה צילום מסך, מלבן, טקסט

התיאור נוצר באופן אוטומטי

גרף 3 MDVP:Fo(Hz) (תדר הבסיס):

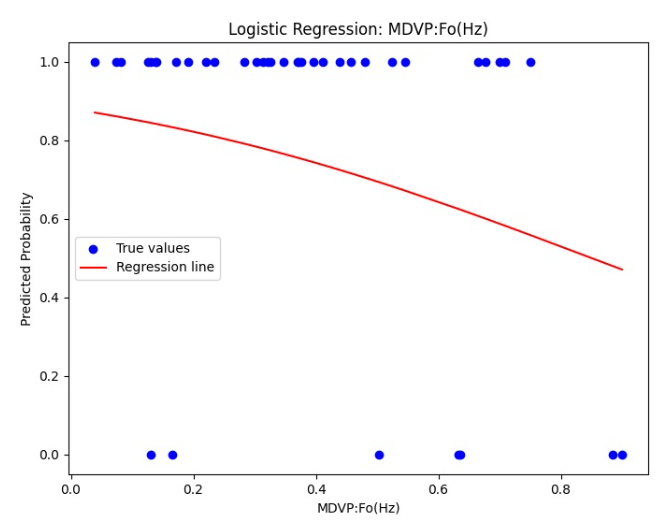
ממוצע הקבוצה הבריאה הוא 0.54

ממוצע הקבוצה של חולי פרקינסון הוא 0.33

בגרף זה קיים גם כן הבדל משמעותי שבו הבריאים מציגים ערכים גבוהים יותר בהשוואה לחולים.

מסקנות כוללות משלושת הגרפים:

* בכל שלושת המדדים (Flo, Fhi, Fo), הקבוצה הבריאה מציגה ערכים ממוצעים גבוהים משמעותית בהשוואה לקבוצת חולי פרקינסון.
* ההבדלים מעידים על ירידה ביציבות ובעוצמת הקול אצל חולי פרקינסון ועשויים לשקף שינויים פיזיולוגיים שנובעים מהמחלה.
* הממצאים משלושת הגרפים תומכים במובהקות ההבדלים ביציבות ובעוצמת הקול בין קבוצת הבריאים לחולי פרקינסון. התוצאות יכולות להוות בסיס למחקרי המשך או לפיתוח שיטות אבחון המבוססות על מדדים קוליים.ראש הטופס

רגרסיה לוגיסטית: ניבוי סטטוס לפי משתנה

הגרף מציג את תוצאות הרגרסיה הלוגיסטית שנעשתה על הנתונים. הנקודות הכחולות מייצגות את הסטטוס האמיתי של המשתתפים, שבו 0 מציין "בריא" ו-1 מציין "חולה". הקו האדום מייצג את ההסתברות החזויה לסטטוס חולה בהתבסס על ערכי. MDVP:Fo(Hz)

השיפוע של הקו מצביע על השפעת MDVP:Fo(Hz) על ההסתברות לסטטוס חולה, ובמקרה זה, ככל שערך MDVP:Fo(Hz) עולה, ההסתברות לחלות יורדת, מה שמעיד על שיפוע שלילי.

מסקנות מרכזיות הן שיש מתאם שלילי בין MDVP:Fo(Hz) לסטטוס החולה, כלומר ככל שערך MDVP:Fo(Hz) יורד, עולה הסיכוי לחזות סטטוס חולה. תחתית הטופס

**בהתבסס על תוצאות המחקר, ניתן לקבוע שההשערה שלנו אכן נכונה. נמצא כי ישנם הבדלים משמעותיים ביציבות ובעוצמת הקול בין חולי פרקינסון לבין האוכלוסייה הבריאה. חולי פרקינסון הציגו יציבות מופחתת בטון ועוצמת קול נמוכה יותר, דבר שמצביע על השפעת המחלה על יכולות הדיבור והתקשורת.**