

## פרויקט גמר פייתון מתקדם

הקדמה :

בפרויקט ישנם שני מאגרי נתונים. האחד מכיל נתונים אודות 42 חולי פרקינסון שנמדדו במדדים שונים 1400 פעמים לאורך תקופה. השני מכיל נתונים אודות 32 נבדקים, 23 מתוכם חולי פרקינסון והשאר בריאים.

בשני המאגרים נתונים אודות מדדים קוליים (HNR, NHR, Shimmers, Jitters וכדו') ובאחד בו מופיעים חולי פרקינסון בלבד ישנם גם ציוני UPDRS של הנבדקים לאורך התקופה.

שאלות המחקר :

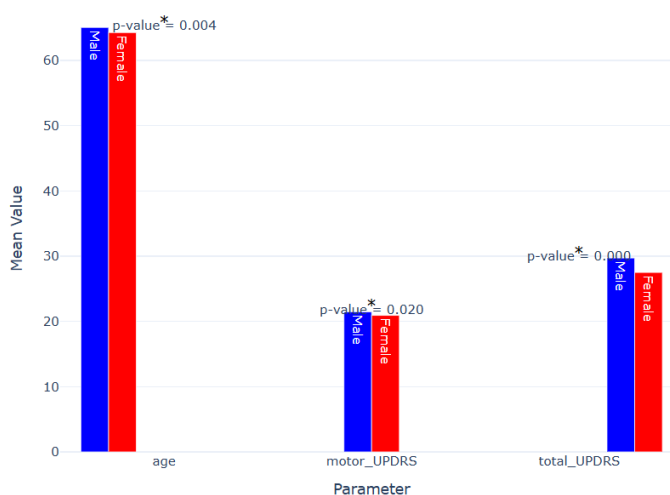
1. אילו הבדלים יש בין גברים ונשים חולי פרקינסון במובנים של מדדי קול והתקדמות המחלה במאגר הנתונים?
2. עד כמה חזקה הקורלציה בין המאפיינים במאגר בחולי פרקינסון? עד כמה חזקה לפי הסטטוס (חולה/בריא)?
3. האם אנחנו יכולים לחזות את התקדמות המחלה (ציוני UPDRS) בהתבסס על המאפיינים השונים בנבדקים ספציפיים ו"בחולה הממוצע"?
4. איך מחלת הפרקינסון מתקדמת עם הזמן?
5. בניית מודל למידת מכונה שיבדיל בין נבדקים בריאים וחולים בעזרת המדדים

שיטות :

בפרויקט בניתי פונקציות שונות בפייתון שבעזרתן עניתי על שאלות המחקר. על מנת לתכלל את כולן וליצור ממשק נוח למשתמש, בניתי ממשק משתמש גרפי שיהפוך את חווית המשתמש לנוחה ופשוטה לתפעול.

תוצאות :

Comparison of Mean Values by Gender (age, motor\_UPDRS, total\_UPDRS)



1.

: Significant Parameters

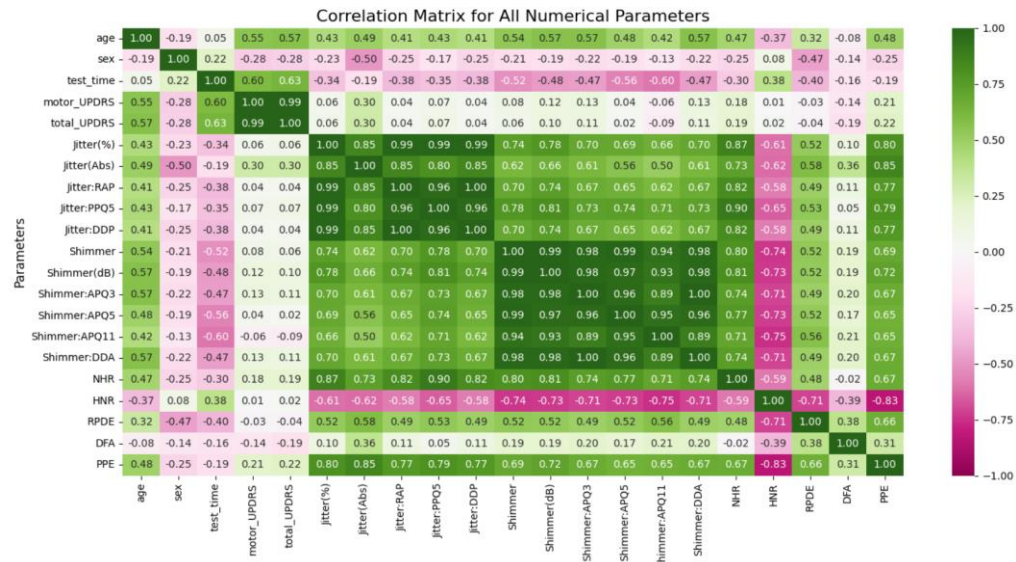
age's p-value = 0.004

motor\_UPDRS's p-value = 0.020

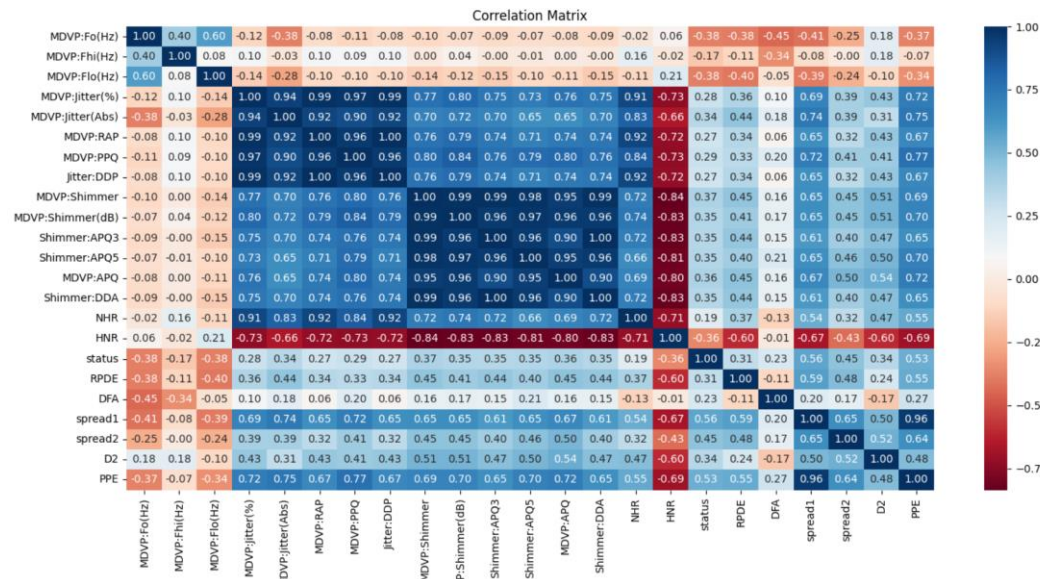
total\_UPDRS's p-value = 0.000

Jitter(%)'s p-value = 0.001  
Jitter(Abs)'s p-value = 0.000  
Jitter: RAP's p-value = 0.000  
Jitter: PPQ5's p-value = 0.000  
Jitter: DDP's p-value = 0.000  
Shimmer's p-value = 0.000  
Shimmer(dB)'s p-value = 0.000  
Shimmer: APQ3's p-value = 0.003  
Shimmer: APQ5's p-value = 0.000  
Shimmer: DDA's p-value = 0.003  
NHR's p-value = 0.000  
RPDE's p-value = 0.000  
DFA's p-value = 0.000  
PPE's p-value = 0.000

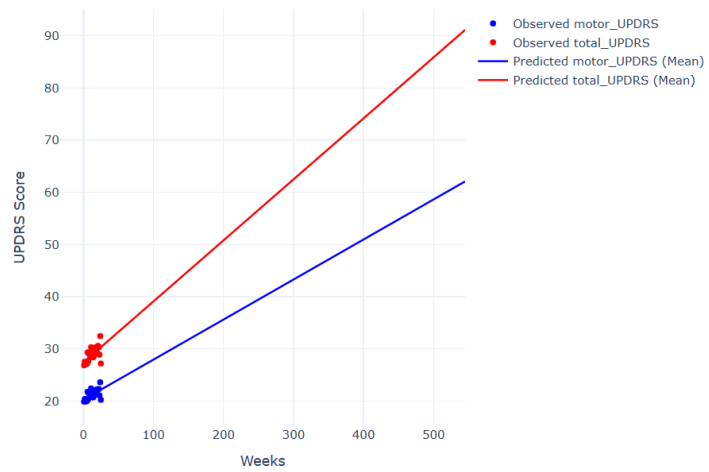
: Non-Significant Parameters  
['test\_time', 'Shimmer: APQ11', 'HNR']



.2

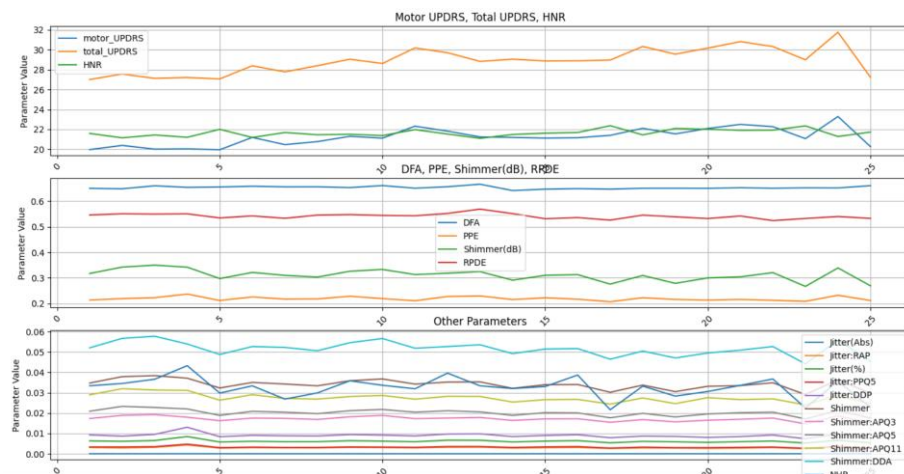
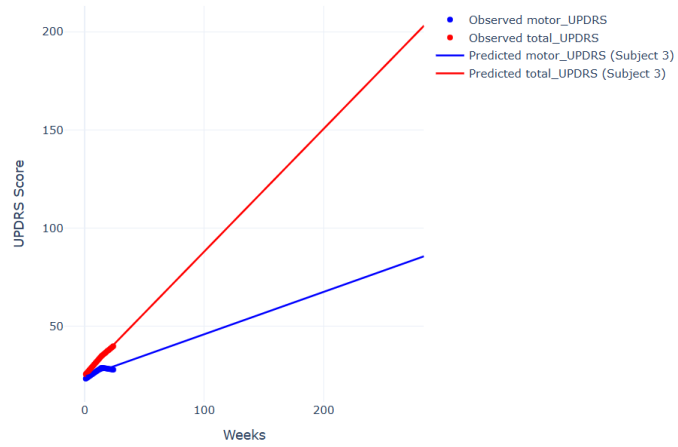


Predicted and Observed Progression of Parkinson's Disease (Mean)

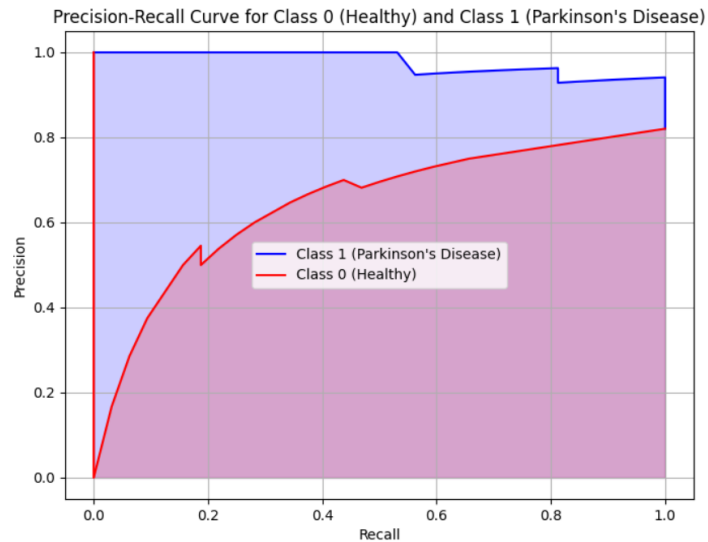


.3

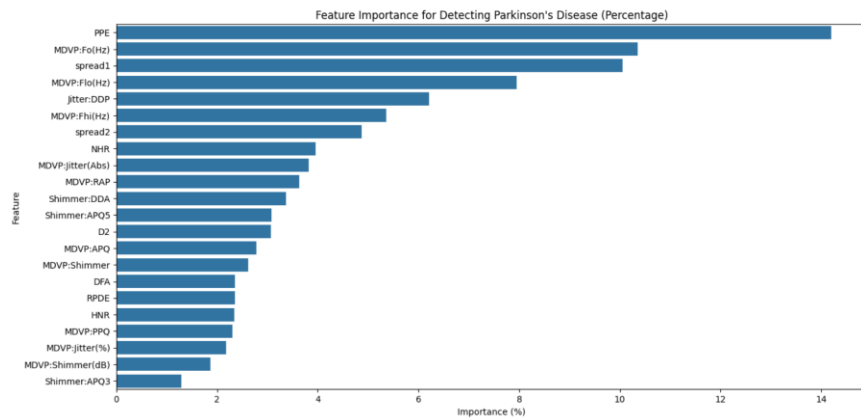
Predicted and Observed Progression of Parkinson's Disease (Subject 3)



.4



.5



Accuracy: 0.9487				
	precision	recall	f1-score	support
0	1.00	0.71	0.83	7
1	0.94	1.00	0.97	32
accuracy			0.95	39
macro avg	0.97	0.86	0.90	39
weighted avg	0.95	0.95	0.95	39

דיון :

מהתוצאות אנחנו לומדים דברים רבים הקשורים לפרקינסון ולהתקדמות המחלה.

ראשית, אנחנו רואים הבדלים מובהקים בין ציוני UPDRS וגילי הגברים והנשים שנבדקו במחקר. אפשר לטעון שההבדל בין ציוני UPDRS מבוססים על ההבדל בגיל, אך בכל אופן ההבדל המשמעותי מצביע על קושי בהשוואת הנתונים של שתי הקבוצות. לכן לא המשכתי להתעמק בנושא – ההבדל יכול להיות מוסבר על ידי גורמים שאינם מין.

בקורלציות של המאגר הראשון אנחנו יכולים לראות כי יש מתאם של 0.6 ו-0.63 בין זמן הבדיקה לציוני UPDRS. דבר זה מצביע על מתאם מסוים בין הזמן להתקדמות המחלה. מעריכה שחלק מהסיבה לכך שהמתאם לא יותר גבוה למרות שאולי היינו מצפים לכך, הוא נטילת תרופות וקבלת טיפולים שונים למחלה, שעלולים לגרום לירידת הציונים של UPDRS.

בקורלציות השניות, ניתן לראות מתאם של 0.53 בין הסטטוס של המחלה לציון הPPE, שהוא הערכה של "fundamental frequency variation" שנמצא במחקרים קודמים כגורם שמופחת בחולי פרקינסון.

במודל שמטרתו לחזות את התקדמות המחלה בחולים שונים אנחנו רואים שהמודל משתנה בהתאם לנבדק אחר, והחיזוי מותאם לתוצאות האישיות של הנבדק. במקרה של "הנבדק הממוצע" אנחנו רואים שבחיזוי של עשור, הנבדק מוערך להגיע לכדי ציון של 90. לעומתו נבדק 3 יעבור את הציון הנ"ל בתוך שנתיים בלבד.

בגרף 4 אנחנו רואים את ההתקדמות השבועית הממוצעת של המאפיינים השונים שנמדדו. ניתן לראות עליה מסוימת בציוני UPDRS, וירידה בציוני הShimmer: APQ11. תוצאות אלו מראות לנו בצורה ויזואלית את התוצאות של הגרף של הקורלציות.

בגרף 5 אנחנו רואים את המודל חיזוי של האם נבדק הוא חולה או בריא. רמת הדיוק של המודל היא 94%, רמה די גבוהה. הגורם המשמעותי ביותר לחיזוי הוא הPPE.

לסיכום, הדבר המשמעותי שניתן לראות הוא שמחלת הפרקינסון מורכבת מגורמים רבים ומשפיעה גם על גורמים קוליים. בפרויקט יצרנו מודל המאפשר חיזוי של התקדמות המחלה בעזרת מעקב ציונים ומאפיינים קוליים של כחצי שנה, ומודל המאפשר אבחון של המחלה בעל דיוק של כ-94%. מודלים אלו יכולים להיות משמעותיים בשימוש עבור אבחון המחלה והערכה של התקדמותה. מודל התקדמות המחלה יוכל גם לשמש עבור בדיקה האם טיפולים מסוימים עובדים טוב לאורך זמן.