



# מחלת הפרקינסון לסוגיה השונים

כאשר רפואה ומדעי המחשב נפגשים...

## תקציר

מחלת הפרקינסון (רטטת בעברית) הינה מחלה כרונית פרוגרסיבית וחשוכת מרפא המתאפיינת ברעד בגפיים, נוקשות שרירים, עיוות בשיווי המשקל אשר מתבטאת בחיי היום יום בעיקר באיטיות פיזית ומחשבתית. את הטיפול במחלת הפרקינסון מנהלים לרוב נוירולוגים אשר מומחים בהפרעות תנועה ופרקים. מתוך מאמרים רבים ניתן לראות כי הטיפול במחלה מתבסס בעיקרו על תרופות שפועלות על גירוי מוחי עמוק (רובן מכילות דופמין). התוצאה היא האטה בהתקדמות המחלה, הפחתת התסמינים הנוירולוגיים ועלייה בתוחלת החיים של הלוקים בה. בעקבות כמות מרובה של חולים במחלת הפרקינסון נכתבו מאמרים רבים על סוגיה השונים של המחלה שנרצה לנצל בכדי לחקור את הקשר, הסיבתיות ואת מגמות התנודות של מחלת הפרקינסון.

איתי גיא

ltgu1010@gmail.com

### שאלת המחקר:

מה הם התנודות בכתיבת מאמרים בנושאי פרקינסון השונים לאורך השנים?

- מענה על שאלת המחקר הנ"ל יתקבל לאחר ניתוח נתונים מעמיק של פלט האפליקציה ע"י ד"ר רות הנדזל.

### תוכן עניינים:

1. מטרת האפליקציה
2. תיאור ותוכנה
3. סיכום ומסקנות

1.

### מטרת האפליקציה:

- מטרתה העיקרית של האפליקציה היא לסנן מתוך מאגר הנתונים של *pubmed* את הנתונים הנכונים שתואמים לשדות הרלוונטיים לשאלת המחקר בכדי שהחוקרת תוכל לראות לפנייה את התפלגות הנתונים באופן ברור במטרה לסייע לה בשלב הצלבת הנתונים וכן בשלב הסקת המסקנות.
- אפליקציה זו התמקדה (בתור התחלה) ב-2 נושאים של מחלת הפרקינסון : גנטיקה ותורת החיסון, כל מאמר התווסף לפלט המערכת רק אם שדה ה-*MajorTopic* שלו זה *parkinson disease* וה-*Qualifier* שלו זה *immunology* או *geneticss*.
- פלט המערכת יוצג בצורת גרף : בגנטיקה של מחלת הפרקינסון ותורת החיסון במחלת הפרקינסון – המציג התפלגות של כמות המאמרים לאורך השנים וקובץ ה-*xlsx* שממנו מוצג גרף ההתפלגות.
- 2.

### תיאור ותוכנה:

תיאור:

### א. הורדת מסד הנתונים:

מסד הנתונים מיוצג באמצעות קובץ *XML* שניתן להוריד באופן הבא :

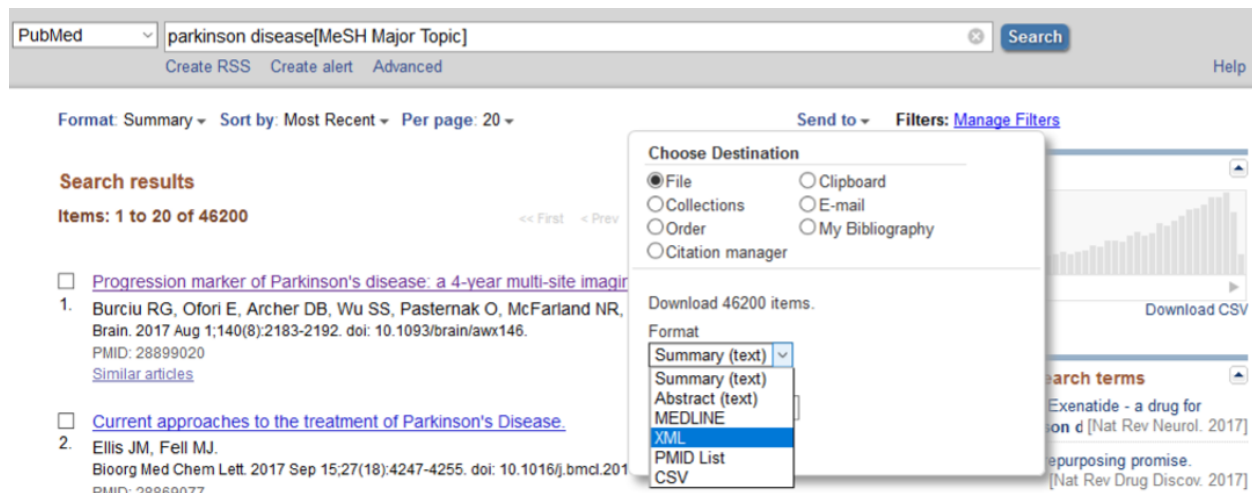
- נווט אל - <https://www.ncbi.nlm.nih.gov/pubmed/advanced>
- תחת *MeSH Major Topic* חפש *parkinson disease* ולחץ על *search*

Builder

MeSH Major Topic	parkinson disease	Show index list
AND	All Fields	Show index list
Search or Add to history		

- הערה – ניתן להוריד כל מסד נתונים אחר ע"י חיפוש של שם אחר.
- בדף שמוצג ישנם רשימה של כל המאמרים שתואמים למילה ולשדה החיפוש שנרצה להוריד בכדי לגשת אליהם באמצעות התוכנה:

- לחץ על *Send to*
- סמן *File*
- בחר ב-*Format* מסוג *XML*
- לחץ כל *Create File*



- הערה – נא לשמור את הקובץ בשם *pubmed\_result.xml* ליד קובץ התוכנה ולא לשנות את שמו כי אחרת התוכנה לא תדע היכן מסד הנתונים.

## ב. שלב הרצת האפליקציה:

בכדי להריץ את האפליקציה ניתן ללחוץ על הקבצים *run1.pyw* [מבצע חיפוש עבור תורת החיסון] או *run2.pyw* [מבצע חיפוש עבור גנטיקה].

קבצים אלו נפרדים מקוד התוכנה בכדי להתקין לפני ריצת התוכנה את כל הספריות שהתוכנה צריכה משפת *python* על מחשב הלקוח בכדי לבצע את פעולתה.

בסיון הפעלת הקבצים הנ"ל יתקבלו 2 פלטים:

- גרף התפלגות של הנושא שנבחר לבצע חיפוש עליו [תורת החיסון או גנטיקה]
- קובץ *results.xlsx* בשם *results.xlsx* המכיל טבלה של הנתונים כאשר:

■ עמודות הטבלה יכילו את השדות הבאים:

$$pmid, title, source, authorName, year, (MeSHTerm, Qualifier)_1, \dots, (MeSHTerm, Qualifier)_n$$

■ כל שורה מכילה מאמר

- קובץ *setupAnalyzer.log* בשם *setupAnalyzer.log* שרושם בתוכו תיעוד של הרצת האפליקציה – מידע תכנותי בלבד המתקבל ע"י הרצת *run1.pyw* וגם *run2.pyw*

- החוקרת מיועדת להשתמש בגרף ההתפלגויות בהמשך המחקר ולשם כך גרף זה ישמר בסיון ריצת האפליקציה באופן אוטומטי כתמונה מפורמט *png*.

- שם תמונת הגרף יבנה באופן הבא – *MeSHTerm\_Date\_Qualifier.png*, לדוגמא :
- עבור גרף שמציג – *MeSHTerm = parkinson disease*, *Qualifier = genetics*, *Date = 17.12.2017*, הוא ישמר בשם *17\_12\_2017\_genetics\_parkinsondisease.png*
- הערה – נא לשים לב שהרצה נוספת של **אותו נושא החיפוש** ידרוס את תמונת הגרף הנוכחית ע"י תמונת גרף חדשה.

#### תוכנה:

האפליקציה נכתבה בשפת *python 3.6.2* ומשתמשת בספריות החיצוניות הבאות :

- *–U pip*
- *numpy*
- *xlsxwriter*
- *progress*
- *–U setuptools*
- *matplotlib*

יתר הספריות הן סטנדרטיות שמובנות בשפת *python 3.6.2*.

#### **השיטה הכללית:**

- *run1* וגם *run2*, לאחר התקנת הספריות החיצוניות במחשב הלקוח, שולחות פרמטר ל-*main* שמייצג את החיפוש שנרצה לבצע – תורת החיסון או גנטיקה
  - ה-*main* מפעיל את *xmlParsing* שאוסף לתוך מבנה נתונים את כל המאמרים שעונים לחיפוש הרצוי תוך בניית השדות הרצויים בטבלת ה-*xlsx*
  - ממיינים את מבנה הנתונים לפי שדה *year* באמצעות *sorted*
  - ממיינים את מבנה הנתונים לפי שדה *pmid* באמצעות *sortEachPmidSortedGroup*
  - מעתיקים את מבנה הנתונים לקובץ *xlsx* באמצעות *insertMatrixToCsv*
  - מבצעים פילטר ששולף את השדות הרצויים להצגת ההתפלגות באמצעות *filterByPair*
  - בונים ומציגים את גרף ההתפלגות באמצעות *showBarChart*
  - שומרים כתמונה ומסיימים
- הערה: בכדי לשנות את הקוד/לשפר/להבין טוב יותר כנס אל קוד המקור בקובץ *PubmedAnalyzer.pyw*

שינוי בפרמטרי התוכנה יתבצע ב-*constructor* של המחלקה:

```
def __init__(self,searchmode):  
    if searchmode == 0:  
        self.mesh = ""  
        self.qualifier = ""  
    elif searchmode == 1:  
        self.mesh = "Parkinson Disease"  
        self.qualifier = "immunology"  
    elif searchmode == 2:  
        self.mesh = "Parkinson Disease"  
        self.qualifier = "genetics"
```

.3

### סיכום ומסקנות:

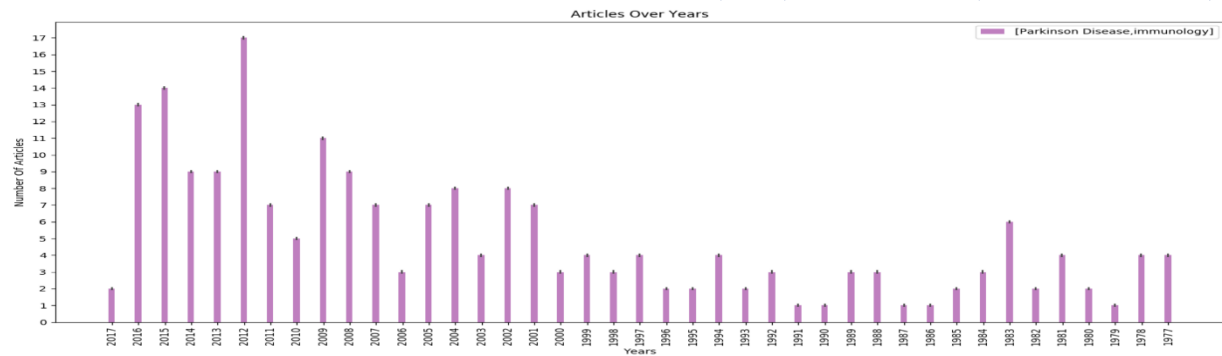
כתיבת האפליקציה חשפה אותי לתחום מרתק ומסתורי בעולם הרפואה שנעזר שיטות במדעי המחשב לצורך מציאת קצה חוט לפתרון.

בנוסף לכך, נחשפתי בצורה עמוקה יותר לשפת תכנות - *python* – המיועדת לניתוח ועיבוד נתונים.

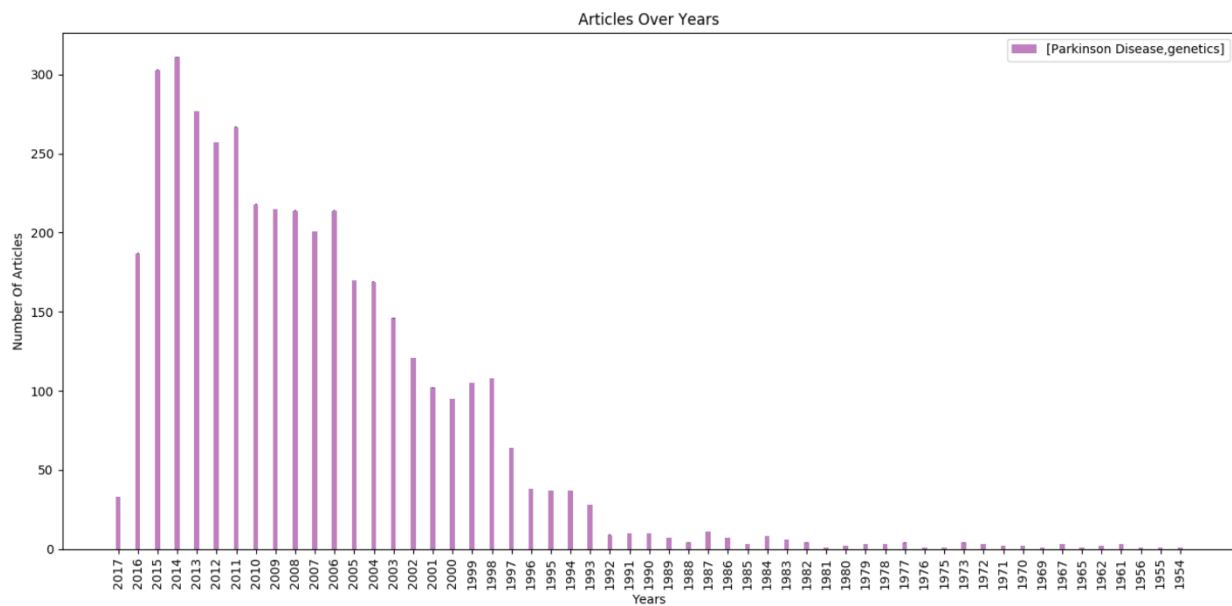
שתי הנקודות שצוינו הנ"ל חיברו אותי באופן אישי לכתיבת הפרויקט.

בכלליות, המוצר הסופי יוכל לעזור בבסיסו למספר רב של חוקרים שירצו לקבל התפלגויות נתונים עבור מאמרים ממסד הנתונים של *pubmed* ללא שינוי או לכל היותר שינוי מינורי בתוכנה [רק בפרמטרים שהוצגו בקוד הנ"ל].

## גרף 1 : תורת החיסון במחלת הפרקינסון



## גרף 2 : גנטיקה של מחלת הפרקינסון



## פלט התוכנה :

הראשון - מתייחס לתורת החיסון במחלת הפרקינסון - כפי שניתן לראות ישנה נקודת מקסימום אחת לצד תנודות לא גדולות בכמות המאמרים שנכתבו.

לעומת זאת, השני - מתייחס לגנטיקה במחלת הפרקינסון - כפי שניתן לראות ישנה עליה חדה בכמות המאמרים החל משנת 1993 ועד היום.

הנתונים הנ"ל יכולים להעיד על המון התרחשויות בנושאים אלו של מחלת הפרקינסון וכמובן שיכולים להיות בסיס להצלבות נתונים אחרות.

תודה רבה על הקריאה!