

אחזור מידע – תרגיל בית 1 – חורף התשפ"ה

מגשים:

צחי בקל 315730176

דניאל ארמגניאן 209146943

יונתן שרר 318317682

ליאור ז'ילגו 316109115

שאלה 1: מודלים לאחזור מידע

.1

Question 4a.a

| | D1 | D2 | D3 | D4 |
|---------------|----|----|----|----|
| breakthrough | 1 | 0 | 0 | 0 |
| drugs | 1 | 1 | 0 | 0 |
| for | 1 | 0 | 1 | 1 |
| schizophrenia | 1 | 1 | 1 | 1 |
| new | 0 | 1 | 1 | 1 |
| approach | 0 | 0 | 1 | 0 |
| treatment | 0 | 0 | 1 | 0 |
| of | 0 | 0 | 1 | 0 |
| hopes | 0 | 0 | 0 | 1 |
| patients | 0 | 0 | 0 | 1 |

Question 4a.b

breakthrough → 1

drugs → 1 2

for → 1 3 4

schizophrenia → 1 2 3 4

new → 2 3 4

approach → 3

treatment → 3

of → 3

hopes → 4

patients → 4

Question 4b.a

schizophrenia → 1 2 3 4q

drugs → 1 2

schizophrenia and drugs: 1 2

Return results: **DOC1, DOC2.**

Question 4b.b

for → 1 3 4

drugs → 1 2

approach → 3

not (drugs or approach): 4

for and not (drugs or approach): 4

Return results: **DOC4.**

Question 4c.a

במקרה זה, נבדוק רק את המסמכים שנמצאים ב Brutus ולא ב Caesar, כלומר נצטרך לעבור פעם אחת על כל אחד מערכי שני הרשימות ולבדוק האם הם קיימים בשניהם או רק ב Brutus. $O(x+y)$ complexity

Question 4c.b

כדי לבצע שלילה ל Caesar, כלומר not Caesar, דורשת מעבר על כל המסמכים הקיימים (נגדיר n). אבל איחוד עם המילה Brutus נדרש גם כן לעבור על כל המסמכים מרשימה זאת, כלומר $O(n+x)$ complexity

| | D1 | D2 | D3 |
|-----------|----|----|----|
| car | 27 | 4 | 24 |
| auto | 3 | 33 | 0 |
| insurance | 0 | 33 | 29 |
| best | 14 | 0 | 17 |
| total | 44 | 70 | 70 |

| | D1 | D2 | D3 |
|--------------|---------------|---------------|---------------|
| car-tf | $27/44=0.613$ | $4/70=0.057$ | $24/70=0.342$ |
| auto-tf | $3/44=0.068$ | $33/70=0.471$ | $0/70=0$ |
| insurance-tf | $0/70=0$ | $33/70=0.471$ | $29/70=0.414$ |
| best-tf | $14/44=0.318$ | $0/70=0$ | $17/70=0.242$ |

| | D1 | D2 | D3 |
|------------------|-----------------------------|-----------------------------|-----------------------------|
| car-tf-idf | $0.613 \times 1.65 = 1.011$ | $0.057 \times 1.65 = 0.094$ | $0.342 \times 1.65 = 0.564$ |
| auto-tf-idf | $0.068 \times 2.08 = 0.141$ | $0.471 \times 2.08 = 0.979$ | $0 \times 2.08 = 0$ |
| insurance-tf-idf | $0 \times 1.62 = 0$ | $0.471 \times 1.62 = 0.763$ | $0.414 \times 1.62 = 0.67$ |
| best-tf-idf | $0.318 \times 1.5 = 0.477$ | $0 \times 1.5 = 0$ | $0.242 \times 1.5 = 0.363$ |

קישור לאתר

האתר של ארגון הבריאות העולמי (WHO - World Health Organization), הוא אתר המתמקד בבריאות הציבור בעולם ומספק מידע על מגוון רב של נושאים בריאותיים. ניתן למצוא בו מידע עדכני אודות התפרצויות של מחלות, מדיניות בריאות ויוזמות בריאות עולמיות. בנוסף, ניתן לאתר בו נתונים, פרסומים מחקריים והנחיות בריאות. מאפשר גישה למידע על מגפות עדכניות, חיפוש מהיר לפי אות ראשונה וקישורים לשותפים לארגון הבריאות העולמי כגון קבוצות מחקר, מימון וגופים האוספים נתונים מרחבי העולם עבורו.

שאלות עליון היינו רוצים לקבל תשובה באתר WHO:

חיפוש מידע על שירותי בריאות מקומיים במדינה מסוימת.
פרטי מידע נדרשים: רשימת רופאים, בתי חולים ומרכזים רפואיים באזור מסוים.

אילו תרופות מתאימות למקרה רפואי ספציפי?
פרטי מידע נדרשים: התאמת תרופות לפי היסטוריה רפואית.

מהם סיכויי ההחלמה של מטופל מסוים עם מחלה ספציפית?
פרטי מידע נדרשים: תחזית אישית של מטופל לפי נתונים כמו גיל, מצב בריאותי כללי, טיפולים נוכחיים.