תרגיל בית 1

מגישים:

שחף ישראל – 209288414

שרה סאלח – 318107448

יובל סאבטו – 318474269

שי אוסמו – 318370285

**שאלה 1:**

סעיף 1:

Question 4a:

a.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **term** | **Doc4** | **Doc3** | **Doc2** | **Doc1** |
| breakthrough | 0 | 0 | 0 | 1 |
| drugs | 0 | 0 | 1 | 1 |
| for | 1 | 1 | 0 | 1 |
| schizophrenia | 1 | 1 | 1 | 1 |
| new | 1 | 1 | 1 | 0 |
| approach | 0 | 1 | 0 | 0 |
| treatment | 0 | 1 | 0 | 0 |
| hopes | 1 | 0 | 0 | 0 |
| patients | 1 | 0 | 0 | 0 |
| of | 0 | 1 | 0 | 0 |

b.

**breakthrough** → [1]

**drugs** → [1, 2]

**for** → [1, 3, 4]

**schizophrenia** → [1, 2, 3, 4]

**new** → [2, 3, 4]

**approach** → [3]

**treatment** → [3]

**hopes** → [4]

**patients** → [4]

**of** → [3]

**Question 4b:**

a.

schizophrenia מופיע במסמכים [1, 2, 3, 4]

drugs מופיע במסמכים [1, 2]

נבצע חיתוך ונקבל [1,2]

b.

שלב 1 – נבצע פעולת איחוד:

drugs מופיע במסמכים [1, 2]

approach מופיע במסמך [3]

נבצע איחוד ונקבל [1,2,3]

שלב 2 – נבצע פעולת שלילה:

כל המסמכים באוסף [1, 2, 3, 4]

נבצע שלילה של [1,2,3] ונקבל [4]

שלב 3 – נבצע פעולת חיתוך:

for מופיע במסמכים [1, 3, 4]

נבצע חיתוך עם [4] ונקבל [4]

לסיכום – מסמך 4

Question 4c:

a.

עלינו לקחת את כל המסמכים שבהם מופיע "Brutus" ולסנן מתוכם את המסמכים שבהם מופיע "Caesar".

נבצע חיתוך מותנה:   
Brutus ∩ (not Caesar)

ננתח את זמן הריצה:

שלילה של רשימת המסמכים של "Caesar" דורשת גישה לכל המסמכים האפשריים במאגר.

לאחר מכן, יש לבצע חיתוך עם "Brutus"

לכן זמן הריצה יהיה O(N+x) כאשר N הוא מספר המסמכים.

b.

עלינו לקחת את כל המסמכים שבהם מופיע "Brutus" ולאחד אותם עם כל המסמכים שבהם לא מופיע "Caesar" .  
Brutus ∪ (not Caesar)

ננתח את זמן הריצה:

* שלילה של "Caesar" (כמו בשאילתה הקודמת) דורשת גישה לכל המסמכים.
* איחוד עם רשימת "Brutus" דורש מעבר על כל המסמכים.

לכן זמן הריצה יהיה O(N).

סעיף 2:

1. חישוב עבור "car":

Doc1: 27⋅1.65=44.55

Doc2: 4⋅1.65=6.64

Doc3: 24⋅1.65=39.624

1. חישוב עבור "auto":

Doc1: 0⋅2.08=0

Doc2: 33⋅2.08=68.6433

Doc3: 0⋅2.08=0

1. חישוב עבור "insurance":

Doc1: 0⋅1.62=0

Doc2: 33⋅1.62=53.4633

Doc3: 29⋅1.62=46.9829

1. חישוב עבור "best":

Doc1: 14⋅1.5=2114

Doc2: 0⋅1.5=0

Doc3: 17⋅1.5=25.517

**שאלה 2:**

**קישור לגיט:**

**https://github.com/shahaf5641/IR---HW1**

**שאלה 3:**

שאילתה 1: מהן המדינות עם שיעור האבטלה הגבוה ביותר לשנת 2023?

פרטי מידע נדרשים:

שם המדינה.

שיעור האבטלה לשנת 2023.

שאילתה 2: מהם מקורות האנרגיה הנפוצים ביותר (כגון פחם, רוח, שמש) בכל מדינה באיחוד האירופי?

פרטי מידע נדרשים:

שם המדינה.

חלוקת מקורות האנרגיה באחוזים.

שאילתה 3: מהן המדינות שבהן אחוז השימוש בתחבורה ציבורית הוא הגבוה ביותר?

פרטי מידע נדרשים:

שם המדינה.

אחוז השימוש בתחבורה ציבורית מתוך כלל צורות התחבורה.