בס"ד

**דו""ח תרגיל 3 בינה מלאכותית:**

**גדי דידי 208064840**

**אליעד סלם 319203162**

חלק א:

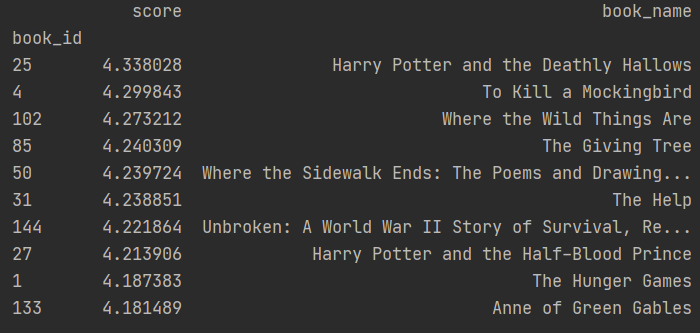
שאלה 1:

הערכים שבחרנו ליצירת מדד הדימיון עבור מערכת המלצה "נונ-פרסונליז" היא

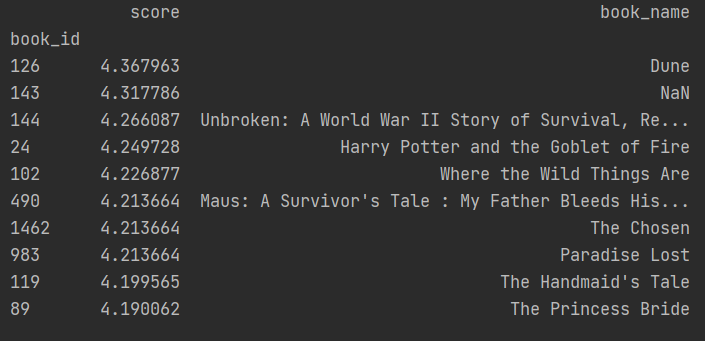
Weighted average rating

בד"כ צריך לקחת עוד מדד מלבד הדירוג ,כמו כמות האנשים **שקראו** (כי ייתכן מצב וספר עם דירוג 5\5 אך רק אדם אחד דירג אותו לעומת ספר 4.3/5 וכמות נכבדת של אנשים דירגה אותו) אך אצלנו לא קיים בדטא את הנתון הנ"ל.

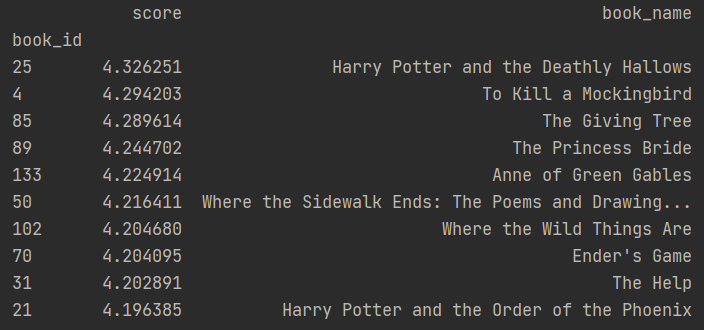
שאלה 2:

K=10 עבור:

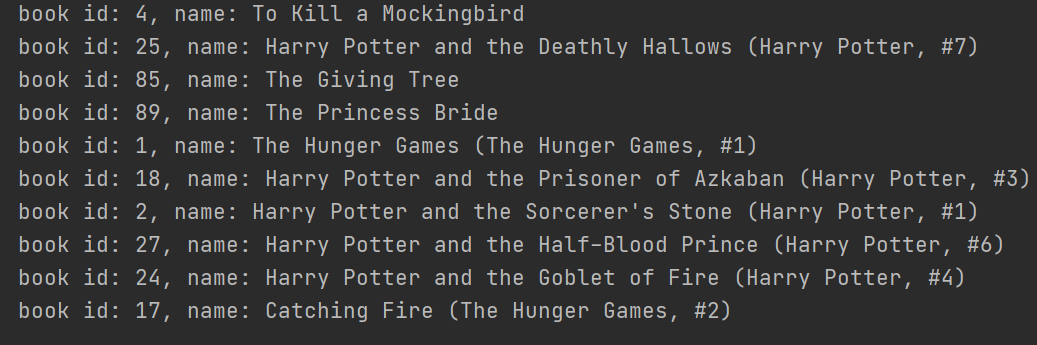
שאלה 3: עבור העיר אוהיו

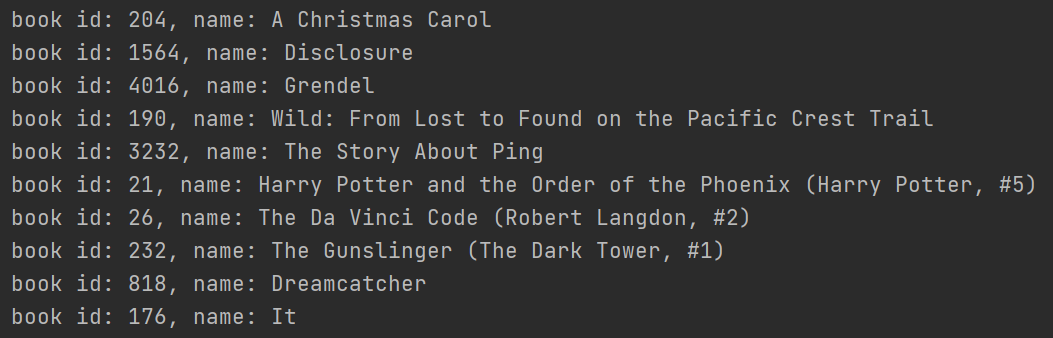
K=10 עבור:

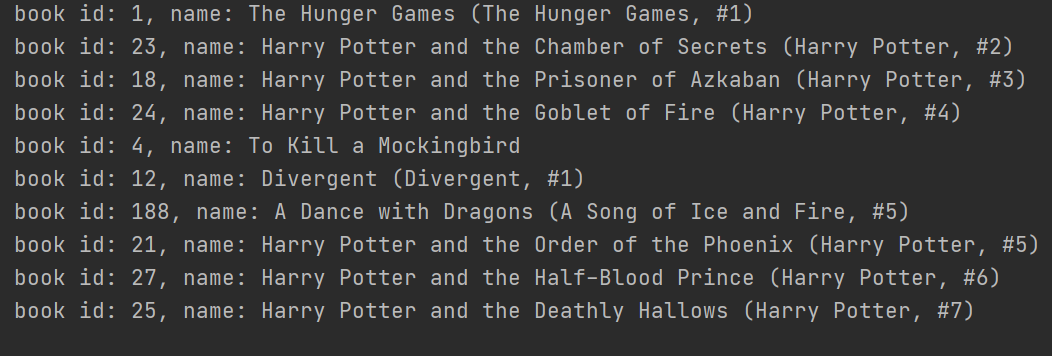
שאלה 4: עבור גיל 28:

K=10 עבור:

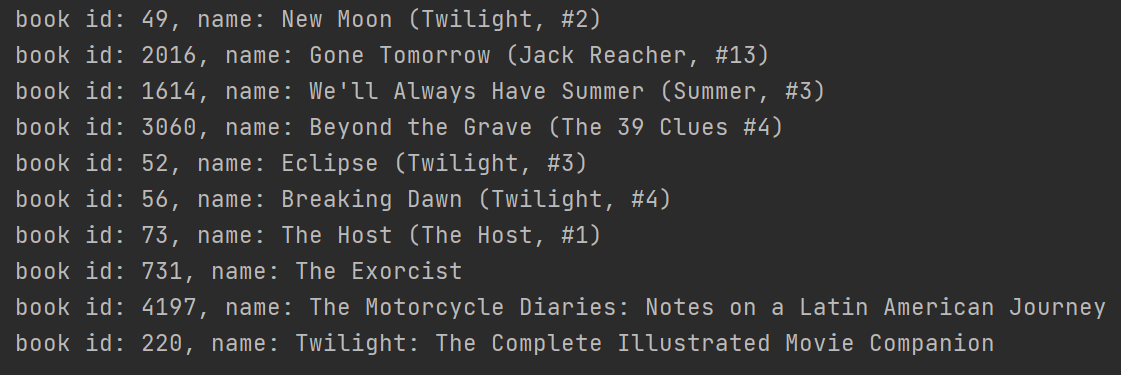
חלק ב':

Cosine top k=10, user=511:

Euclidean top k=10, user=511:

Jaccard top k=10, user=511:

חלק ג':

Name book: Twilight, top k =10

חלק ד':

|  |  |  |  |
| --- | --- | --- | --- |
|  | **precision\_k** | **ARHR** | **RMSE** |
| **cosine** | 0.088 | 0.3354 | 0.9009 |
| **Euclidean** | 0.008 | 0.033 | 0.9191 |
| **jaccard** | 0.088 | 0.3342 | 0.9049 |

שאלה 12:

-RMSE

המדד מחשב את המרחק בין החיזוי לתוצאה האמיתית ככל שהתוצאה תהיה גדולה כך רמת הדיוק תהיה נמוכה.

ניתן לראות על פי הנתונים שיש דיוק נמוך בין טבלת החיזוי שלנו לבין הדירוגים עצמם של המשתמשים מאחר והמרחק גבוה..

-percision@k מחזירה לנו ממוצע עבור כמות הפגיעות בהעדפות המשתמש ביחס למה שהבאנו לו. מדד זה מתייחס לדירוגים על סרטים שדורגו 4-5.

- ARHR חישבנו הוספנו עבור כל פגיעה בהעדפת משתמש אמיתית מתוך העדפות שהמערכת נתנה לו התייחסות למיקום שלו בסדר העדיפויות ולבסוף חילקנו בכמות היוזרים .

מאחר וב- RMSE קיבלנו ערכים גבוהים ביותר לכן ב- precision@k ,ARHR קיבלנו ערכים יחסית נמוכים מפאת חוסר דיוק של מטריצת הדימיון.

ומצד ההבדלים בין precision@k ו ARHRזה נגרם וגורם החילוק ,באחד חילקנו במיקום ההמלצה ובשני בסך כל ההמלצות לכן ARHR ייתן ערכים גבוהים יותר .