



מכללת כנפי רוח, קרית נוער ירושלים

סמל מוסד 140129

ספר פרויקט לעבודת גמר

י"ד הנדסאי תוכנה (שאלון 714918)

בנושא:

מערכת המלצה למשחקי ווידיאו

מגיש: ינון יאיר זוהר

ת"ז: 213882442

מנחה: אילן פרץ

תשפ"ד 2024

תוכן עניינים

הצעת הפרויקט שאושרה ע"י משרד החינוך	5
1.1 פרטי מגיש ההצעה	5
1.2 פרטי הפרויקט	5
1.3 הגדרת הבעיה האלגוריתמית	5
1.4 5 רקע תיאורטי בתחום הפרויקט	
1.5 הליכים עיקריים	5
1.6 מסד הנתונים	5
1.7 5 פרטים פורמליים	
מבוא	5
2.1 הרקע לפרויקט	5
2.2 תהליך המחקר	5
2.3 סקירת ספרות	5
2.4 אתגרים מרכזיים במחקר	5
2.4.1 התמודדות עם הבעיה	6
2.4.2 הסיבות לבחירת הנושא	6
2.4.3 על איזה צורך הפרויקט עונה ?	7
2.4.4 הצגת הפתרונות שנבחרו לבעיה	7
מטרות ויעדים	
אתגרים בבניית המערכת	8
הגדרת מדדי הצלחה למערכת	9
רקע תיאורטי וספרות מקצועית	9
9 תיאור מצב קיים	
3 ניתוח חלופות מערכת	13
4 תיאור החלופה הנבחרת ונימוקים לבחירתה	14
5 איפיון המערכת המוצעת	14
6 ניתוח הדרישות מהמערכת	14
7 מודולי המערכת	14
8 אפיון פונקצונאלי וביצועים עקריים	14
9 אילוצי המערכת	14
10 שגיאה! הסימניה אינה מוגדרת. תיאור תהליכי אבטחת מידע במערכת (רק למי שיש)	
11.1 שגיאה! הסימניה אינה מוגדרת. תיאור התקיפה / ההגנה	

שגיאה! הסימניה אינה מוגדרת.....	11.2
למידת מכונה(איסוף נתונים)	14
ארכיטקטורת המערכת (ע"פ מודל 3 השכבות)	12.1
שכבת הלוגיקה	12.1.1
שכבת הנתונים	12.1.2
שכבת התצוגה	12.1.3
12	12.2
תיכון מפורט של רכיבי המערכת	12.2
17 מודל שרת/לקוח	12.3
תיאור מסד הנתונים	12.4
שגיאה! הסימניה אינה מוגדרת.....	12.5
תיאור פרוטוקולי תקשורת	12.6
שגיאה! הסימניה אינה מוגדרת.....	12.7
ChatGPT , REST API שירותים חיצוניים (אם יש כגון :	12.8
חלופות לתיכון המערכת	12.9
ניתוח תרחישים וזרימת המידע	20
Use-Case Diagram תרשימי תרחיש	13.1
הצגת מקרי שימוש	13.1.1
תיאור מקרי שימוש עיקריים	13.1.2
Sequence Diagram תרשימי רצף	13.2
13	13.2.1
login(התחברות)	13.2.1
חיפוש משחק	13.2.2
גלישה ברשימת ההמלצות	13.2.3
פעולה 4	13.2.4
תיאור מסכים וממשק משתמש	20
תרשים היררכיית המסכים	14.1
תיאור המסכים	14.2
תיאור התוכנה	20
סביבת עבודה	15.1
14	15.2
שפות תכנות	15.2
תיאור המודולים והמחלקות	15.3
עץ המודולים	15.3.1
15	15.3.2
Class Diagram תרשים מחלקות	15.3.2
תיאור מחלקות מפורט	15.3.3
מבנה נתונים בשימוש	15.4

15.5.	קוד התוכנית לפעולות/פונקציות/אלגוריתמים עיקריות וחשובים	21
	מדריך למשתמש	21
	בדיקות והערכה	21
	מסקנות.....	21
	פיתוחים עתידיים	21
	בבליוגרפיה.....	22
16.		
17.		
18.		
19.		
20.		

הצעת הפרויקט שאושרה ע"י משרד החינוך

מבוא

1.

2.1. הרקע לפרויקט

2ני משחק משחקי מחשב מאז שאני ילד, ואחד הדברים שכל גיימר מתקשה איתו זה להחליט על איזה משחק לשחק. השוק כיום מוצף באלפי משחקים שונים בסגנונות וז'אנרים רבים. ומשחקים חדשים יוצאים בקצב שקשה לעקוב אחריו. הרבה גיימרים כיום מתקשים למצוא משחקים חדשים מעבר למה שהם כבר מכירים דבר שמביא לחוסר גיוון. לכן החלטתי לבנות אתר שיאפשר לגיימרים להכניס את הרגלי המשחק שלהם ולקבל המלצות על משחקים שיתכן שיהנו מהם

2.2. תהליך המחקר

המחקר שלי עסק בעיקר במערכות ההמלצה השונות והחסרונות של כל אחד מהם והשימושים השונים בהם. במהלך המחקר למדי על האלגוריתמים השונים והצורה שבא כל אחד מהם עובד. לאחר שלמדתי על הטכנולוגיות השונות הגעתי למערכת מסוג collaborative filtering שנראו כאילו הם יתאימו למשימה שלי (פירוט בסעיף הבאה)

בשלב החקר של הפרויקט עסקתי במערכות המלצה. עיקר המערכת שלי היא מערכת ההמלצה אז התחלתי לחקור את הנושא.

התחלתי בלחקור על הפתרונות הקיימים לבעיה. כדי להבין יותר לעומק את התהליך הבסיסי של יצירת ההמלצות, הדברים הדרושים ליצירה של מערכת מסוג זה והבנה של האתגרים שיכולים לעלות במהלך הבניה שלה בנוסף חקרתי יותר לעומק את הטכנולוגיות השונות להקמת אתר שנלמדו בכיתה כדי לוודא התאמה בינם לבין המערכת הנבחרת

2.3. סקירת ספרות

התייחסות למקורות המובאים למטה והסבר קצר על מה כל אחת עוסקת

2.4. אתגרים מרכזיים במחקר

האתגר המרכזי בבניית הפרויקט היה פיתוח אלגוריתם מתאים, קיימות מספר רב של שיטות לבניית מערכות המלצה, לכל אחד מהם יתרונות וחסרונות ומקרים בהם עם עדיפות על פני שיטת אחרות. בחירה של האלגוריתם הטוב ביותר למקרה שימוש שלי, היה דבר קריטי, שישפיע על הרבה פרטים של הפרויקט בהמשך הפיתוח. הדבר הצריך מחקר מעמיק ומחשבה רבה, הבנה של השיטות האחרות ושיקולי דעת שונים. דבר נוסף שבו מצאתי קושי היה בחירת הפרמטר שעליו יחושבו ההמלצות. דבר זה משפיע רבות על המערכת שכן בחירת פרמטר לא אמין תביא לקושי במתן המלצות מדויקות ואיכותיות למשתמש, בנוסף הבחירה הראשונה של פרמטר (דירוג משתמש) הוכחה את עצמה כלא אמינה. דבר שהוביל להליכה אחורה בפיתוח והצריך מחקר נוסף.

שנית בחירת הכלים לשימוש פיתוח האתר הוא גם כן חלק חשוב, יש להשקיע מחשבה רבה בבחירת כלים שמתאימים לא רק למשימה אלה גם למתכנת, בחירת כלים לא מתאימים יביא לאי יעילות קושי והארכת זמן הפיתוח.

קישור המודל לאתר. עוד אתגר בבניית הפרויקט היה ההטמעה של האלגוריתם במערכת, המערכות השונות צריכות לתקשר באופן יעיל עם האלגוריתם ונצרך פיתוח רב של מערכות תומכות כדי לשלב את האלגוריתם עם שאר המערכת.

2.4.1. התמודדות עם הבעיה

הבעיה המרכזית שהצריכה התמודדות הייתה בניית מערכת המלצות שמייצרת המלצות איכותיות בזמן סביר, שתאפשר יצירת המלצות לכל המשתמשים תוך שימוש בכמה שפחות משאבים.

ראשית יש לפתח אלגוריתם על פי השיטה הנבחרת, ולבצע בדיקות של האלגוריתם תוך שיפור ביצועים ופתירת תקלות.

לאחר מכן יש לשלב את האלגוריתם באופן יעיל בתוך המערכת של הפרויקט, דבר שמצריך התאמה לא רק של האלגוריתם אלה גם של המערכת ובניית תוכנות תומכות לאלגוריתם המרכזי שיאפשרו לו לקבל את הנתונים הדרושים לו.

לאחר מכן יש לבצע בדיקות נוספות של המערכת עם האלגוריתם עם המערכת, למצוא בעיות של אי תאימות ולפתור אותן.

2.4.2. הסיבות לבחירת הנושא

מערכות המלצות נמצאות בכל מקום, הם מעצבות את החוויה שלנו ברשת מבלי שנשים לב בכלל. בין אם מדובר בהצעת סרט חדש בפלטפורמות צפיה בסרטים או בהצעת מוצרים באתרי קניות, המערכות הללו לומדות את ההתנהגות שלנו ואת העדפות שלנו. לכן בפרויקט שלי, רציתי ללמוד על המערכות האלו. איך הן עובדות. ומכיוון שאני חובב משחקי מחשב מאז שהייתי ילד, רציתי לקשר בין השניים ולבנות מערכת המלצה בנושא משחקי מחשב

2.4.3. על איזה צורך הפרויקט עונה ?

הפרויקט נותן מענה לאנשים שלא יכולים להשקיע זמן בלחפש ולחקור משחקים שונים כדי לבחור מה לשחק, האתר נותן מענה של מערכת המלצות שחולשת על כל הפלטפורמות המכירה השונות שרבים מהם לא מציעים מערכת המלצות בכלל. ההמלצות מותאמות אישית למשתמש במקום להציע רק את מה שפופולרי עכשיו ככה שהם מגוונות יותר, וכן סיכוי יותר טוב שהלקוח יהיה מרוצה מהם (מאחר והם מורכבות מהעדפות שלו) בנוסף האתר יהיה נגיש מכל מכשיר בעל דפדפן ככה שהלקוח לא יצטרך להשתמש במחשב/קונסולה כדי לראות את ההמלצות שלו ולעדכן את הנתונים שלו במערכת, כך שיהיה אפשר להשתמש בו גם שאין לך גישה למערכת המשחקים המועדפת עליך

2.4.4. הצגת הפתרונות שנבחרו לבעיה

הפתרון לבעיה הוא אלגוריתם המלצה

אלגוריתם המלצה הוא אלגוריתם שמשתמש בנתונים על המשתמשים כדי ללמוד את העדפות שלהם ולהמליץ על דברים חדשים
קיימים כמה סוגי מערכות המלצה שניתן לחלק לארבע סוגים

- מערכות מבוססות תוכן - מתבססות על תוכן או תיאור הפריטים. המערכות בקטגוריה זו מתחלקות לשלושה סוגים מבוססות על תוכן פרופיל משתמש ותיאור פריטים, מבוססות על נתוני פרופיל המשתמש (גיל, מיקום גאוגרפי ועוד) ומבוססות על היסטוריית הגלישה של המשתמש
- מערכות מבוססות משוב – מחולקות לשני סוגים Explicit Feedback וImplicit Feedback
- מערכות סינון שיתופי- מערכות שמסתמכות על הדימיון בין העדפות של משתמשים כדי למצוא משתמשים בעל טעם דומה ולהמליץ בהסתמך על מידע זה.
- מערכות משולבות (HYBRIDS) - מערכות אלו משלבות בין שיטות השונות על מנת לתת המלצות מדויקות יותר.

מטרות ויעדים

לפרוייקט כמה יעדים מרכזיים

- **חויית משתמש - פיתוח מערכת משתמש נגישה וקלה לשימוש שתאפשר ללקוח להשתמש במערכת באופן מהיר ונוח**
- **תמיכה נרחבת - לאפשר תמיכה של כמה שיותר פלטפורמות משחקים, תוך עדכון לתמיכה ופלטפורמות שונות ומשחקים חדשים**
- **איסוף מידע ושמירה על פרטיות - יצירת מאגר נרחב של פרטי משתמש לשיפור החוויה תוך שמירה על פרטיות המשתמש**
- **שימוש במערכות המלצה מתקדמות - שימוש במערכות המלצה מתקדמות ליצירת המלצות מדויקות ומותאמות אישית לכל משתמש**

אתגרים בבניית המערכת

- **איסוף מידע – למערכת נדרש מידע רב כדי לתפקד, איסוף המידע תוך שמירה על פרטיות הלקוח הוא חלק קריטי בבניית המערכת**
- **עיבוד המידע – על מנת שהאלגוריתם יעבוד בצורה הטובה ביותר המידע נדרש לעבור תהליך של עיבוד, ניקוי המידע הוא הליך הכרחי לדיוק האלגוריתם**
- **בניית האתר – בניית האתר מצריכה עבודה בתחומים רבים, שימוש בכלים חדשים שאין לי ניסיון קודם איתם בניה של המערכת הפנימית וכן מערכת הממשק משתמש דורשת הבנה מלאה בתחומים האלו.**
- **בניית האלגוריתם והטמעתו במערכת – בניית אלגוריתם יעיל ומדויקת שיבצע את המשימה בזמן סביר והטמעתו המערכת**

הגדרת מדדי הצלחה למערכת

- שמירת המידע בצורה יעילה ואמינה
- שימוש בהתחברות לשמירת המידע של כל משתמש
- ממשק משתמש ברור יעיל ונגיש
- 5. נתינת המלצות מותאמות אישית לכל משתמש על בסיס המידע שלו

רקע תיאורטי וספרות מקצועית

מערכות המלצה הם סוג של אלגוריתמים שמנסה להבין למה אנשים אוהבים דברים מסוימים על מנת לנחש מה אנשים יאהבו, יש מספר שימושים למערכות מסוג זה כמו המלצה של מוצרים בחנויות אינטרנט, או מחקר של מוצר לצורך שיפור שלו, והוא הליבה של הפרויקט שלי.

6. בנושא הראשון שחקרתי היה מערכות המלצה, מהם? איך הם פועלות ואיך משתמשים בהם. לאחר מחקר קצר הבנתי שיש שתי אפשרויות להמלצת דברים לשחקנים. המלצות חכמות והמלצות "טיפשות"

האפשרות הראשונה היא פשוטה ואינה דורשת שימוש באלגוריתמים מסובכים או מידע נרחב על המשתמש, והיא בעצם להמליץ דברים פשוטים כמו דברים חדשים שיצאו לאחרונה, מוצרים בהנחה, או מוצרים פופולאריים.

השיטה השנייה מערבת שימוש באלגוריתמים מתקדמים ובמידע על המוצרים והמשתמשים כדי לתת המלצות חכמות יותר שמותאמות אישית למשתמש.

החלטתי שאני רוצה ליצור מערכת הסוג השני בעלת המלצות חכמות מותאמות למשתמש כדי לספק המלצות איכותיות יותר לכל משתמש

החומר הראשון שקראתי בנושא היה (ראה מקור אחד) Recommender Systems — A Complete Guide to Machine Learning Models

המאמר הציג את הנושא של מערכות המלצה חכמות, דיבר על המטרה של מערכות מסוג ושימושים השונים בשבילם, בנוסף המאמר הציג כמה מושגים בסיסיים בנושא כמו כן את סוגי המערכות השונות שקיימות, איך כל אחת עובדת והיתרונות והחסרונות שלה

המאמר הסביר על שימוש בExplicit Feedback לעומת Implicit Feedback (משוב מפורש לעומת משוב מרומז)

כמו כן על הסוגים השונים של המערכות. מערכות מבוססות תוכן, מערכות מסוג סינון שיתופי ומערכות היברידיות

עכשיו נסביר בקצרה על כל סוג

מערכות מבוססות תוכן משתמשות במידע על המוצר (התוכן שברצוננו להמליץ) כדי להמליץ אותו למשתמשים שונים,

מערכות מבוססות משוב - מערכות אלו תלויות בשיתוף פעולה מצד המשתמש. קטגוריה זו מחולקת לשיטות המבוססות על דירוג פריטים ושיטות המבוססות על דירוג קשרים בין פריטים

מערכות סינון שיתופי- מערכות שעובדות על ההיגיון שמשתמשים שהסכימו בעבר יסכימו בעתיד

מערכות משולבות (HYBRIDS) - מערכות אלו משלבות בין השיטות השונות על מנת לתת המלצות מדויקות יותר

מערכות מסינון שיתופי נראו כדבר המושלם לפרויקט שלי משום הנטייה של גיימרים להתחלק לקבוצות מעריצים בעלי דעות דומות,

לכן החלטתי לחקור לעומק יותר את הנושא של collaborative filtering כדי לראות את ההתאמה שלו לפרויקט אז קראתי את הערך בוויקיפדיה על הנושא (ראה מקור 3)

שם העמקתי יותר על השיטות השונות לביצוע של collaborative filtering הקשיים ששימוש במערכת מכל סוג יכול לגרום, והשימוש המתאים לכל אחד

לאחר סקירה של המערכות השונות הגעתי למסקנה ש matrix factorization נראה מתאים למקרה שימוש שלי וכן יהיה אפשרי למימוש במסגרת הפרויקט אז המשכתי את המחקר בנושא המחקר כלל מאמרים רבים וכמו כן סרטונים בנושא, (מקורות 3-7)

אחד הדברים המרכזיים שעליו דובר היה בחירת ה metric (המידע שעל פיו יבוצעו ההמלצות) של המערכת

תחילה רציתי להשתמש במשוב מפורש על ידי דירוג של כל משתמש על המשחקים שהוא שיחק. שיטה זו זאת התגלתה כלא יעילה למקרה שלי. על פי מקור 8 לגיימרים יש קשיים בדירוג דברים באופן הוגן וקיים קושי במציאת קורלציה בין דירוג שניתן על ידי השחקן לרמת ההנאה שלו מהמשחק, על פי נתון מהמחקר, במקרים מסוימים הסיכוי ששחקן ייתן דירוג טוב למשחק יורד ככל שהוא משחק בו יותר. ישנם שחקנים ששיחקו במשחק אלפי שעות, ובכל זאת מדרגים אותו נמוך, דבר זה לא מאפשר שימוש בדירוג של משתמשים כנתון לחישוב ההמלצה. מכיוון שהוא אינו מדויק ובלתי צפוי, ניתן היה לראות שיש צורך במציאת metric אחר

לאחר מחקר נוסף ומחשבה על metrics אפשריים נתקלתי בדף מחקר (ראה מקור 9) שחוקר את השימוש בזמן משחק כנתון לבניית המלצות הדף דיבר בחיוב על השימוש בזמן משחק לבניית המלצות ועל פי נתונים שלו המערכת ביצעה באופן שווה או טוב יותר ממערכות דומות. לכן החלטתי גם כן להשתמש בזמן משחק כmetrics למערכת ההמלצות שלי

בנ

תיאור מצב קיים

על פי סקירה של האפשרויות בקיימות כרגע נראה שאין מערכת שמציעה את מה שאני מנסה לספק על המערכת שלי

על פי סקר שערך yougov.com (ראה מקור 10) הרבה שחקנים לא יודעים מהיכן לגלות משחקים חדשים

יותר מ 15% לא יודעים מהיכן הם מגלים משחקים ויותר מ 25% לא משתמשים באף אחד מהדרכים שהוצעו

תחום רכישת המשחקים עבר שינוי מאוד נרחב בעשור האחרון, המעבר ממשחקים פיזיים (דיסקים, קסטות וכדומה) למשחקים דיגיטליים גרם לכך שהאפשרויות נרחבות מתמיד, הגישה לכל משחק דרך האינטרנט נתן לשחקנים אפשרות לשחק משחקים שלא הייתה להם גישה אליה לפני כן.

כמו כן תעשיית הגיימינג גדלה מאוד בשנים האחרונות ושווה יותר מ 500 ביליון דולר. זה אומר שחברות חדשות מצטרפות כל הזמן והאפשרויות רק הולכות וגדלות. אם לפני עשר שנים היו כמה חנויות בודדות שמכרו משחקים כיום יש עשרות פלטפורמות שונות שנועדו למכור אותם, דבר שגורם לקושי משמעותי בבחירה של משחקים

דבר נוסף שגורם לקושי במציאת משחקים היא החלוקה של פלטפורמות המכירה בין החברות והבלעדיות של חלק מהפלטפורמות על משחקים מסוימים.

שתעשיית המשחקים התחילה לעבור לאונליין הייתה פלטפורמה אחת מרכזית שמכרה משחקים, steam. הדבר הקל מאוד על השחקנים וכן על המפתחים שכן הכל היה מרוכז במקום אחד. אבל החברות האחרות החלו לארח פלטפורמות משלהם כדי להימנע מהאחוזים steam לוקחת מכל מכירה, והחלו למכור את המשחקים שלהם בפלטפורמות ייחודיות משלהם, דבר

שאולי גרם לתחרות אבל גם לריבוי פלטפורמות מכירה, דבר שגורם לאי נוחות מיותרת לשחקנים עד כדי כך שהם לא ירכשו משחק אם אינו נמכר steam למשל משתמש הגיב בדיון בנושא (ראה מקור 11) "האם תסרב לרכוש משחק אם אינו מוצע מכירה ב steam

"Yup, I refused to buy starcraft. I definitely would have bought if it were available ".on Steam

תרגום: "כן, אני מסרב לרכוש את סטארקראפט. בוודאות הייתי רוכש אותו אם היה מוצע למכירה ב STEAM

משתמש אחר הגיב
"Never; I am, however, sometimes tempted to not to buy a game if I know it has "to be registered on steam

תרגום: "לא. אבל לפעמים אני נמנע מלקנות משחקים אם אני יודע שהם לא מוצעים steam"

דבר זה מראה שיש פירוז וצורך בשוק למערכת אחת שבא שחקנים יוכלו לנהל את אוסף המשחקים שלהם
פלטפורמות קיימות להמלצת משחקים
קיימות כמה פלטפורמות שנותנות מענה להמלצת משחקים
בראשונה קיימות פלטפורמות המכירה, יש עשרות פלטפורמות כאלה אז נפרט את הבולטים שבהם

- Steam
- EA (לשעבר origin)
- Epic games
- Gog

בנוסף קיימים שירותים חיצוניים למציאת משחקים הבולטים שבהם

- [/https://ggapp.io](https://ggapp.io)
- [/https://www.completionator.com](https://www.completionator.com)
- <https://rawg.io>

בנוסף קיימים אתרים יותר ממוקדים כגון

- Gameslikefind
- quanticfoundry

ניתוח חלופות מערכת

פלטפורמות המכירה - steam, EA, epic games ועוד

השימוש בפלטפורמות המכירה לצורך קבלת המלצות היא הדבר ההגיוני והטבעי לעשות, הם מכילים מידע על משחקים רבים ונותנים למשתמש לבצע את הרכישה וההורדה של המשחק מאותו מקום. למערכות אלה כמה חסרונות עיקריים. הם משתמשות במערכות מהסוג הבסיסי ביותר. לאף אחת מהם אין אלגוריתמים שמתאימים אישית את ההמלצות למשתמש, אלה כולם משתמשות בשיטה הבסיסית של המלצה של מה שפופולארי כרגע, משחקים חדשים, ומשחקים בהנחה בנוסף היותם פלטפורמות מכירה יש לה אינטרסים אחרים מעבר להנאה של השחקן מהמשחק (כדון קידום ממון) בנוסף כל פלטפורמה יכולה להציע רק את המשחקים שהיא מציעה למכירה, דבר המביא לגיוון מופחת שירותים חיצוניים למציאת משחקים ומעקב אחרי משחקים קיימים

אתרים חיצוניים לביצוע מעקב – ROGW, completionator, ggapp

ישנם כמה אתרים שמציעים שירותי מעקב אחרי משחקים אתרים אלה מעולים לאנשים שרוצים לעקוב אחרי המשחקים ששיחקו בצורה נוחה וכן לעקוב אחרי משחקים שהם מתכננים לשחק בעתיד על פי מחקר שלי אף אחד מהן לא מציע אלגוריתם מותאם אישית למשתמש וגם הם משתמשים במערכת בסיסית של המלצת משחקים פופולריים והמלצה של משחקים מסדרת שהמשתמש כבר שיחק הם אכן שיפור מהמערכות של פלטפורמות המכירה. שכן הם מתקדמות יותר והם יכולות להציע משחקים ללא הגבלה של פלטפורמה, אבל מערכת ההמלצה שלהם אינם מתאימות את עצמם למשתמש

• אתרים ממוקדים – quanticfoundry quanticfoundry

אתרים אלו מאפשרים למשתמש לענות על שאלון כדי לקבל המלצות על משחקים חדשים השאלון יכול לכלול רשימה של משחקים שנהנתה מהם או שאלות אישיות על התכונות שלך כשחקן אתרים אלו הם פתרון טוב למי שרוצה לקבל המלצות באופן מהיר אך יש להם כמה חסרונות הם מוגבלות במידע שהן יכולות לקבל, במקרים רבים הם לא מאפשרים הכנסה של מידע רב מספיק, ולא תמיד השאלון יכיל את כל הפרמטרים הנדרשים לביצוע המלצות מדויקות

משימוש שלי בכמה מהם נראה שרובם עובדים על מודלים פשוטים שמציעים המלצות בעיקר על בסיס נתונים יבשים כגון ז'אנר או מפתח מסוים, והם נוטים להציע משחקים מאותה הסדרה בנוסף הם לא מציעים התחברות ככה שמרגע שהמשתמש עוזב את הדף כל המידע שהוא הכניס אבד והוא יידרש להכניס אותו מחדש בפעם הבאה, דבר שמונע מהאתר להכיר את המשתמש באופן מעמיק

תיאור החלופה הנבחרת ונימוקים לבחירתה

הפתרון שאני בחרתי הוא יצירת אתר שיאפשר למשתמשים לקבל המלצות על מה לשחק מחוץ להגבלות של פלטפורמות המכירה השונות. האתר יאפשר למשתמשים להזין את נתוני המשחק שלהם ולקבל על פיהם המלצות מותאמות אישית להעדפות שלהם, המערכת תאפשר התחברות ⁹ ככה שהיא תזכור את נתוני המשתמש בנוסף המערכת תציע מקום מרוכז בשביל המשתמש לעקוב אחרי המשחקים שהוא שיחק מכל מקום מבלי להיכנס לכל פלטפורמה בנפרד, הבחירה במערכת זו היא משום שהיא עונה על כל הדרישות לפתרון. ותיתן מענה טוב לשחקנים ואפיון אפקטיביות של מערכות דומות למוצרים בקטגוריות כגון <https://www.imdb.com> לסרטים או <https://myanimelist.net> לסדרות וספרים ממוצא יפני

איפיון המערכת המוצעת

10.1. ניתוח הדרישות מהמערכת

10.2. מודולי המערכת ¹⁰

10.3. אפיון פונקציונלי וביצועים עיקריים

10.4. אילוצי המערכת

למידת מכונה (איסוף נתונים)

אף על פי שהמערכת שלי אינה נחשבת בינה מלאכותית, יש לה הרבה במשותף עם מערכות מהסוג הזה

לצורך בדיקות של האלגוריתם, היה צורך בנתונים על משתמשים קיימים, השימוש בנתונים אמיתיים הוא קריטי להערכת המודל. שכן מטרת האלגוריתם היא למצוא קשרים חבואים במידע, בנוסף הנתונים האלו שימשו לפתירת בעיות של jumpstart (מקרה שבו לאלגוריתם אין מספיק מידע כדי לעבוד)

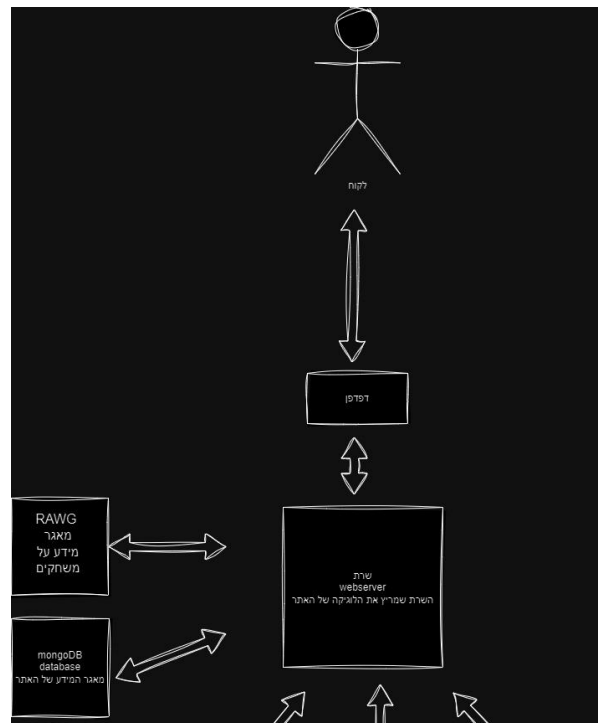
על מנת לאסוף את הנתונים פניתי לאתר Kaggle שהוא אתר המציע סטי נתונים בתחומים רבים לשימוש בבינה מלאכותית ועוד, לאחר חיפוש קצר נתקלתי במאגר מידע של שעות משחק של שחקנים ב־FLTPורמת מכירת המשחקים STEAM. המאגר מכיל מידע של אלפי משמשים הכולל את המשחקים שהם רכשו\שיחקו ואת כמות הזמן שבילו בכל משחק. המידע הזה שימש אותי לבדיקת המערכת וישמש למערכת כמידע התחלתי לפני שהמערכת תאסוף מספיק מידע על משתמשים

תיכון המערכת

12.1. ארכיטקטורת המערכת (ע"פ מודל 3 השכבות)

להלן תיאור ארכיטקטורת הפרויקט

12.1 תרשים למטה ניתן לראות את המבנה הבסיסי של המערכת



12.1.1. שכבת הלוגיקה

בשכבה הזאת נמצא השרת, השרת אחראי להפעיל את האתר ולממש את הפעולות שלו. השרת הוא הגב של כל המערכת והכל עובר דרכו, הוא מפעיל את האתר, שולט בנתונים, מממש פעולות ועוד.

השרת שלי הוא שרת TOMCAT מסוג HTTP Webserver

12.1.2. שכבת הנתונים

במערכת שלי יש שני מערכות מאגרי נתונים מרכזיות הראשונה היא מאגר הנתונים האישי של הפרויקט, השרת משמש כזכרון של השרת והוא משמש את השרת לאחסון מידע כגון שמות משתמש, סיסמאות, מידע על משחקים ששוחקו ועוד השרת שלי הוא שרת מסוג שיושב בענן והוא ניתן כשירות על ידי MongoDB בנוסף לשרת הזה, קיים עוד בסיס נתונים שהשרת המרכזי מתקשר איתו השרת הוא שרת רשת שמאורח על ידי חברה חיצונית, השרת מכיל מידע על משחקי וידאו, ומשמש את השרת של הפרויקט כדי לקבל מידע על משחקים שקיימים בשוק, השרת ניגש אליו בעזרת restAPI ומקבל את התשובות בצורת Jason

12.1.3. שכבת התצוגה

שכבת התצוגה היא השכבה שאחרית להציג את המידע למשתמש, ולאפשר לו לבצע אינטרקציה ופעולות עם השרת, שכבת התצוגה שלי היא אתר אינטרנטי שיכול להיפתח דרך דפדפן, בעזרת שכבת התצוגה יכול הלקוח לבצע פעולות עם המערכת והיא הדרך היחידה לשימוש המערכת על ידי לקוח

12.2. תיוון מפורט של רכיבי המערכת

שכבת התצוגה (ממשק משתמש)

שכבת התצוגה של המערכת בנויה על גבי אתר אינטרנטי, הקמת אתר נעשתה על ידי שימוש ב vaadin שהיא פלטפורמת פיתוח אתרים בשפת java שמאפשרת שימוש ברכיבים כגון כפתורים תיבות טקסט ועוד על מנת לבנות אתר בעל עיצוב אחיד בצורה נוחה ומהירה ללא צורך בלמידת שפות לבניית אתרים כמו java script,

שכבת הלוגיקה

שכבת הלוגיקה משתמש בשירות tomcat הוא שרת רשת המריץ Servlets (במינוח רשמי : Servlet container) בקוד פתוח. שעליו יכול לרוץ קוד Java. הוא גם משמש כשרת ברירת המחדל בפיתוחי Java על בסיס Spring (עליו נפרט בהמשך). הוא מאפשר הקמה מהירה של שרתי רשת תוך שימוש בקוד java

תשתית התוכנה ספרינג (באנגלית : Spring Framework ; להלן : ספרינג) היא שלד תוכנה בקוד פתוח ומנגנון להיפוך שליטה (Inversion of Control) עבור הפלטפורמה של ג'אווה.

ניתן להשתמש ביכולות הליבה של ספרינג בכל יישום Java, אבל קיימות גם הרחבות לבניית יישומי רשת על גבי פלטפורמת ה-Java EE. מבוצע בעיקר באמצעות הזרקת תלויות (dependency injection).

שכבת הנתונים

שכבת הנתונים ממומשת על ידי MongoDB
MongoDB הוא מסד נתונים בקטגוריית NoSQL,

בסיס הנתונים נשען על מבנה של מסמך (Document-Oriented Database) בניגוד למסדי נתונים טבלאיים (כמו SQL Server, Oracle ו MySQL) העובדים מעל טבלאות מקושרות. במבנה מסמכים במבנה JSON

12.3. מודל שרת/לקוח**צד השרת**

לבניית צד השרת עשיתי שימוש בspring boot שעושה שימוש בtomcat ביחד הם מטפלים בריבוי משתמשים על ידי הקמת פורטים, וכן קבלה ושליחה של בקשות http, על ידי שימוש בהם ניתן לבנות שרת רשת בצורה קלה מבלי לכתוב את הצד התקשורתי של השרת מאפס

צד הלקוח

בצד הלקוח נעשה שימוש בדפדפן, השימוש בדפדפן חוסך מהמתכנת ליצור אסליקציה מאפס וכן מאפשר הרצה על מכשירים רבים מבלי כתיבת צד לקוח נוסף,

12.4. תיאור מסד הנתונים

מסד הנתונים הוא MongoDB שאינו טבלאי (noSQL) זה אומר שהוא אינו עובד בטכנולוגיה של טבלאות אלא בטכנולוגיה של אוספים, הוא שומר את המידע במבנה של JSON

תיאור המסמכים המאוכסנים במאגר

משתמש

User Collection	
_id	ObjectId
un	string
pw	string
playedGames	array of Game
interestingGames	array of Game

מאחסן מידע של משתמש לכל משתמש יש

- _Id מזהה של האובייקט הניתן על ידי mongoDB
- UN שם המשתמש של הלקוח
- סיסמא הסיסמא שמשתמש להתחברות הלקוח
- Played games מערך של משחקים שהמשתמש שיחק (יפורט בהמשך)
- Interesting games מערך של משחקים שהמשתמש מתעניין בהם

משחק

Game Collection	
_id	ObjectId
name	string
releaseDate	string
Rating	double
platforms	array of string
coverImageLink	string
playTime	int

מבנה המשמש לאחסון משחקים במאגר המידע
כל משחק מכיל

- *Id _ מזהה של האובייקט הניתן על ידי MongoDB
- Name שם המשחק
- releaseDate תאריך יציאת המשחק
- Rating דירוג המשחק
- coverImageLink קישור לתמונת הכריכה של המשחק
- playtime זמן שהמשתמש בילה במשחק

12.5. תיאור פרוטוקולי תקשורת

HTTP

הוא פרוטוקול תקשורת שנועד להעברת דפי HTML ואובייקטים שהם מכילים (כמו תמונות, קובצי קול, סרטונים וכו') ברשת האינטרנט וברשתות אינטראנט

TCP

פרוטוקול בתקשורת נתונים הפועל בשכבות התעבורה של מודל ה-OSI ובמודל ה-TCP/IP, ומבטיח העברה אמינה של נתונים בין שתי תחנות ברשת מחשבים באמצעות יצירת חיבור מקושר (Connection Oriented).

12.6. שירותים חיצוניים

בפרויקט נעשה שימוש בשרת חיצוני המכיל מידע על משחקי מחשב הקיימים בשוק. השרת נקרא RAWG והוא משמש את הפרויקט על מנת לחסוך את מטלת אחסון המידע על המשחקים מהשרת. השירות ש RAGW מספק מאפשר לי לקבל מידע על משחקים מבלי לאחסן את המידע אצלי בצורה מהירה ויעילה, בין השאר השירות משמש על מנת לבדוק עם משחק קיים, לקבל מידע כגון שם, תאריך יציאה, פלטפורמות נתמכות, ותמונות כריכה של המשחקים, שימוש בשרת מסוג זה הוא חיוני לפרויקט לשם הצגת המידע למשתמש וכן לשימוש המידע על המערכות הפנימיות. הגישה לשרת נעשית על ידי rest API שהוא תבנית תקשורת מבוססת http בין שרתים והחזרה של המידע בפורמט JASON השימוש ב REST API מפשט את השימוש בשרת ומאפשר תקשורת חלקה ופשוטה בין השירות לשרת הפרויקט

12.7. חלופות לתיכון המערכת

ניתוח תרחישים וזרימת המידע

13.1. תרשימי תרחיש Use-Case Diagram

13.2. תרשימי רצף Sequence Diagram

13

תיאור מסכים וממשק משתמש

תיאור התוכנה

14.15.1. סביבת עבודה

סביבת עבודה שהשתמשתי בה בפרויקט היא Microsoft visual studio code (vsCode) היא סביבת פיתוח שמאפשרת תמיכה במספר רב של שפות ותכונות על ידי שימוש בplugins (תוספים). השימוש בתוספים ב vscode אפשר לי לשלוט בחלקים השונים של הפרויקט בצורה נוחה בעזרת תוספים לרכיבים השונים, למשל תוסף של MongoDB (מסד הנתונים הנבחר של הפרויקט) מאפשר שליטה במסד בצורה נוחה ויעילה מבלי לעזוב את סביבת הפיתוח, בנוסף החיבור של הכל במקום אחד מאפשר ניפוי שגיאות קל יותר ויעיל יותר

15.2. שפות תכנות

שפת התכנות ששומשה לפרוייקט היא java. Java היא שפת תכנות מונחת עצמים. בחרתי להשתמש בשפה זאת בגלל שיש לי כבר ניסיון קודם עם השפה וזה הקל עלי בכתובת הפרוייקט, בנוסף היא יכולה לרוץ על סוגי מכשירים רבים. בנוסף היות השפה מונחת עצמים הופכת את הניהול של מאגר הנתונים לפשוט יותר

15.3. תיאור המודולים והמחלקות

15.3.1. עץ המודולים

15.3.2. תרשימי מחלקות Class Diagram

15.3.3. תיאור מחלקות מפורט

15.4. מבנה נתונים בשימוש

בפרויקט נעשה שימוש בכמה מבני נתונים

מטריצות

מטריצה היא מערך דו ממדי, הפרויקט עושה שימוש במטריצות בשביל לאכסן את פרטי המשתמשים בצורה שתאפשר לאלגוריתם לעשות חישובים בצורה יעילה ומהירה. דוגמא לשימוש

15.5. קוד התוכנית לפעולות/פונקציות/אלגוריתמים עיקריות וחשובים

לתאר 4-5 פעולות/פונקציות עיקריות/חשובות, ועל כל אחד : להציג הקוד שלו, תיעוד טענת כניסה ויציאה (פרמטרים וערך מוחזר), מה מבצע?, אופן פעולתו בצורה מפורטת, סיבוכיות זמן ריצה...

- פעולה/פונקציה 1
- פעולה/פונקציה 2
- פעולה/פונקציה 3
- פעולה/פונקציה 4

מדריך למשתמש**בדיקות והערכה****16.****מסקנות**

17. לפני שהתחלתי בבניית הפרויקט לא השקעתי את הזמן הדרוש במחקר מקדים על הנושא. חשבתי שאוכל ללמוד תוך כדי עשייה ולאסוף רק את מה שאני צריך. במהלך בניית הפרויקט נתקלתי בהרבה בעיות שהחזירו אותי אחורה לחקור נושאים מחדש, דבר שהיה יכול להימנע ולחסוך זמן בבניית הפרויקט.

18.

הפרויקט בנוסף למדתי על החשיבות של סדר וארגון של המשימות הנדרשות. ברגע שאתה מסדר בראש את הבעיה ומחלק אותה למשימות קטנות הרבה יותר קל לבצע התקדמות משמעותית מבלי לחזור אחורה כדי לתקן ולשנות דברים שלא נעשו כמו שצריך.

פיתוחים עתידיים

- הוספת אלגוריתמים נוספים ושילוב התוצאות שלהם לשיפור איכות ההמלצות
- חיבור לחשבונות בפלטפורמות המכירה לשליפת המידע על זמן המשחק באופן אוטומטי

19.

בבליוגרפיה

1. <https://towardsdatascience.com/recommender-systems-a-complete-guide-to-machine-learning-models-96d3f94ea748>
2. <https://medium.com/@khang.pham.exact/what-are-recommendation-systems-6bb5036042db>
3. https://en.wikipedia.org/wiki/Collaborative_filtering
4. [https://en.wikipedia.org/wiki/Matrix_factorization_\(recommender_systems\)](https://en.wikipedia.org/wiki/Matrix_factorization_(recommender_systems))
5. <https://towardsdatascience.com/recommendation-system-matrix-factorization-d61978660b4b>
6. <https://www.mygreatlearning.com/blog/matrix-factorization-explained>
7. <https://www.youtube.com/watch?v=ZspR5PZemcs>
8. <https://medium.com/@jonduke90/analyzing-steam-reviews-and-users-data-7a4ff3c5ce1a>
9. https://www.researchgate.net/publication/330249306_Estimated_Rating_Based_on_Hours_Played_for_Video_Game_Recommendation
10. <https://business.yougov.com/content/47794-how-gamers-across-6-key-markets-discover-new-video-games>
11. <https://business.yougov.com/content/47794-how-gamers-across-6-key-markets-discover-new-video-games>
12. <https://www.giantbomb.com/forums/general-discussion-30/poll-did-you-not-buy-a-game-because-not-on-steam-w-512176>

