

פרויקט- מערכות נבונות ואינטרקטיביות

מגישות:

אביה שער	ת.ז 305252355
טל אברג'ל	ת.ז 316541101
הילה בן חמו	ת.ז 312141732
תאריך:	29/06/2020

מבוא:

כיום ישנם הרבה שיעורי ספורט משותפים כאשר חלקם מתמקדים במוזיקה (למשל זומבה). בתור צרכניות של עולם זה נתקלנו לא פעם במצבים בהם תוכן השיעור, כלומר השירים שהושמעו לא היו לטעמינו.

את הפרויקט שלנו נעשה בהשראת המאמרים:

"DJ-MC: A Reinforcement-Learning Agent for Music Playlist Recommendation" [4] בו

יוצרים מערכת ליצירת פלייליסט אישי ובודקים את יעילותה.

והמאמר **"EduRank: A Collaborative Filtering Approach to Personalization in E-learning"** [1]

שבעזרת אלגוריתמים מבחירה חברתית מדרג שאלות קיימות לפי רמת קושי ולומד כיצד לדרג את השאלות החדשות שטרם נפתרו.

בעקבות חשיפתינו למאמרים בנושא וכהרחבה למאמר הראשון [4], בחרנו לבנות מערכת שעל סמך הפלייליסטים האישיים תיצור פלייליסט משותף אשר ישפר את חוויית השיעור עבור המשתתפים ויגביר את המחויבות של המשתתפים בשיעור.

תוצאות צפויות

בעזרת ממשק למשתמש שניצור נאסוף דירוגים אישיים של משתמשים כך שיהווה לנו כ"פלייליסטים אישיים". ממימוש אלגוריתמי הבחירה החברתית שהצגנו מעלה, אנו צופות לקבל שני פלייליסטים שיתופיים סופיים שיהוו דירוג המבטא את ההעדפות האישיות של המתאמנים בצורה בה מרביתם יתחברו ויאהבו פלייליסט זה. דירוגים אישיים אלה של הפלייליסטים מהווים דירוגים חלקיים מאחר וכל "מתאמן" מדרג מספר מסוים של שירים מתוך כלל המאגר שניתן. לכן נצפה שהדירוג הסופי שיתקבל יהווה את החציון של הדירוגים החלקיים, כלומר- הדירוג החלקי בעל המרחק הכולל המינימלי לדירוגי הקלט.

להערכתנו, הפלייליסטים שנקבל יהיו דומים מאחר ולכולם המטרה להוות דירוג סופי מתכלל של העדפות המשתמשים. בנוסף, באבולוציה על הפלייליסטים שנוצרו, נצפה לקבל ממוצע שביעות רצון גבוה (בין 7-9 כאשר הדירוג הוא בין 1-10) עם סטיית תקן נמוכה.

סקירה ספרותית.

ראשית, ניתן לראות את הקשר בין מוסיקה לבין התחושות שמועברות באימון ספורט בכלל וריקוד בפרט. ממאמר שעוסק בקשר זה [8] ראינו כי ספורטאים שנשאלו על תחושותיהם בעת אימון ספורט המלווה במוסיקה פירטו על התחושות החיוביות ברמה גבוהה מאוד ובתדירות רבה, לעומת תחושות שליליות שכמעט ולא הוזכרו כלל. תחושות כמו אושר, עירנות, ביטחון, הרפיה הוזכרו במידה שכיחה מאוד בקרב הספורטאים בעיקר בזמנים שבהם הם מבצעים פעילויות הכנה כמו חימום ואימוני הכנה לתחרויות. מגביר את התחושות החיוביות ומקטין את התחושות השליליות.

בשנים האחרונות הגישה למוסיקה התפתחה טכנולוגית לפי המאמר [4]. אם פעם יכולנו לשמוע מוסיקה דרך הרדיו או ערוץ טלוויזיה, כיום עם אתרי האינטרנט נוצר מושג חדש הנקרא רדיו אינטרנטי שבו משדרים או מוציאים לאור שירים ולמאזין יש את הבחירה אם לשמוע את השיר פעם אחת, לחזור לשמוע אותו שוב או לא לשמוע אותו לעולם. כתוצאה מהחשיפה לכמויות אדירות של שירים בצורה הזו, החוקרים רצו לסייע למאזינים בכך שיבנו מערכת שתמליץ להם על שירים שיאהבו ושמומלץ להם לשמוע באמצעות פידבקים של משתמשים קודמים בטכנולוגיה שקראו לה "rich get richer".

כמו שהחוקרים בונים למאזינים מערכת שממליצה אילו שירים מתאימים למאזין לשמוע, גם אנו נרצה לייצר סוג של מערכת המלצה (בפרויקט שלנו נתייחס לפלייליסט כמערכת המלצה). במערכת של המאמר מתבססים גם על דעת ההמונים על ידי שימוש בפידבקים של מאזינים קודמים וזה מחזק את ההשערה שלנו שדעה של הרוב נחשבת מאוד בקרב הציבור ולכן יהיו יותר אנשים שיהיו שבעי רצון מהפלייליסט שניצור. כמו כן, ישנה התמקדות הולכת וגוברת במחקר מערכות ממליצות אוטומטיות. מערכות המלצות מוזיקה משמשות תחום בולט עבור מערכות כאלה, הן מבחינה אקדמית והן מבחינה מסחרית. במאמר [3] מוצגת מסגרת חדשה להמלצת פלייליסט של מוסיקה מותאמת אישית בשם DJ-MC. במסגרת זו ביצעו שימוש בMDP לטובת בחירת רצף שירים לפלייליסט, תוך התמקדות במאפייני השמע של שירים. DJ-MC בעל יתרון בכך שהוא מסוגל לייצר רשימות השמעה המותאמות למאזין ואינן מתחשבות בסיווגים מסורתיים המבוססים על ז'אנר, תקופה וכו'.

לפי המאמר [5] המתאר מערכת המלצה של שירים שפועלת בצורת פילטר המתאים את עצמו למשתמש באופן ישיר, ישנה חשיבות רבה למי מבצע את השיר יותר מאשר התוכן של המילים. המערכת מתבססת על שני סוגי מערכות המלצה: הראשון על דימיון של אמנים והשני על דימיון אקוסטי של תוכן השירים. החוקרים מאמינים כי כאשר מאזינים רואים בצורה ויזואלית את השמות של המילים והאמנים זה משפיע על המאזינים בצורה לא אובייקטיבית. לכן המערכת רוצה להתגבר על ההטיות הללו ולדרג שירים לפי הרמיזות שיש בתוך השירים. גם במאמר זה [5] ניתן לראות כי קיים בסיס לצורך של המלצות לשירים לפי ז'אנרים (אומנים) או סגנון מוזיקלי (תוכן אקוסטי) וכי ישנה חשיבות רבה למי מבצע את השיר יותר מאשר התוכן של המילים. נרצה לבדוק האם מאזינים ידרגו במקומות זהים שירים של אותו אומן מדעה קדומה עליו. במאמר [6] מתבצעת השוואה בין הביצועים של מספר אלגוריתמים ליצירת פלייליסט. ההערכה המתבצעת במאמר זה מתבצעת בכמה דרכים: הערכה אנושית, שבה המשתמשים מתבקשים לענות על שאלון אשר מסייע לאלגוריתם ליצור את הפלייליסט עבורם ולאחר מכן כשיאזינו לו יתבקשו לדרג אותו בסולם של 1-10 ובנוסף למסור את כמות השירים שהשביעו את רצונם וקלעו במידה הגבוהה ביותר לטעמם. ההערכה השנייה היא לכידות סמנטית, כאן

משתמשים בגישות רבות למדידה ודירוג כל שיר ביחס לשירים אחרים בצורה כמותית למשל לכידות שמתבצעת על ידי ספירת תדרים של התרחשויות משותפת של שירים מאת אותו אומן או אנטרופיה של התפלגות ז'אנרים ברשימת ההשמעה. צורת הערכה נוספת היא חיזוי רצף, דרך ישירה יותר להערכה אוטומטית שנובעת מהגדרה פורמלית של פלייליסט כבעיית חיזוי.

בהינתן שאילתה מסוימת שמייצגת את העדפות המשתמש האלגוריתם צריך לחזות מהו השיר הבא שיושמע. לאחר מכן נעריך את האלגוריתם לפי ביצועי החיזוי שלו תחת מוסכמות שהניחו מראש. מתוך מאמר זה נרצה לשאוב רעיונות לביצוע הערכה של איכות וטיב האלגוריתמים שניצור שכן במאמר הם הצליחו להוכיח את נכונות האבלואציה שביצעו.

עלתה בפנינו השאלה כיצד פלייליסט חברתי יכול לתרום למתאמנים? במאמר העוסק בהשפעת בחירה חברתית על קבלת החלטות [9] היווכחנו כי בכל סיטואציה שקיימת אשר דורשת לקבל החלטה (במקרה שלנו אפשר להשליך את ההחלטה על הבחירה של השירים האופטימלית ביותר של המדריכה לפני השיעור על מנת להביא למקסימום הנאה של המתאמנים) בחירה חברתית תורמת בצורה משמעותית לקלות שבה המדריכה יכולה להחליט אילו שירים ייכנסו לפלייליסט מתוך מאגר עצום ומגוון של שירים אשר לעיתים יכול לבלבל ולהטעות את המדריכה. לכן לפי מאמר זה ניתן להסיק כי המערכת אותה ניצור תתרום להנאה של המשתתפים. מה שנותר לנו לדעת הוא בכמה השתפרה מידת ההנאה של המתאמנים.

עולם המוסיקה הינו רחב ובעל מגוון רחב של שירים וסגנונות שכל אחד מתחבר באופן שונה לכן בפרויקט ניתן למשתמשים לבחור מתוך מאגר מסוים כאשר ההנחה היא שלא כולם ידרגו את אותם השירים ולכן יש להתייחס לדירוג חלקי. במאמר [2] בתחום הבחירה החברתית, מוצגת תחילה הבעיה הנובעת מדירוג חלקי הרווח במציאות להשוואות בין אלמנטים שונים. דירוג חלקי מתקבל למשל כאשר מדרגים i מתוך n אלמנטים כך ששאר האלמנטים מקבלים דירוג זהה $i+1$ ונוצר שוויון או כאשר מדובר בשני פרמטרים שאינם ניתנים להשוואה. על מנת להשוות בין דירוגים חלקיים הוצעו מספר דרכי מדידות, שניים מתוכם שמתייחסים אליהם במאמר הם הכללות של דרך מדידה למרחק בין דירוגים מלאים הידוע בשם קנדל טאו. בעזרת חישוב, לא רק שניתן לבצע השוואה בין הדירוגים אלא אף ניתן לבצע אגרציה לדירוגים החלקיים לטובת דירוג יחיד. בפרויקט שלנו נעזר באלגוריתמים המתוארים במאמר ע"מ למצוא הסכמה משותפת כאשר כל אדם מדרג את רשימת השירים שלנו ובעזרת האלגוריתמים נגיע לרשימה אחת משותפת על כל המדרגים.

במאמר נוסף העוסק בדירוגים [1] מתואר אלגוריתם בשם EduRank אשר מטרתו לייצר דירוג אישי על קושי שאלות. עבור כל סטודנט, EduRank מעריך עד כמה סטודנטים אחרים דומים לו ואז משלב את הדירוג של התלמידים הדומים כדי ליצור דירוג עבור כולם. ישנם שני נהלים עיקריים באלגוריתם: חישוב מדד הדמיון של התלמיד ויצירת דירוג קושי על בסיס דירוג של משתמשים דומים. בפרויקט שלנו נבחר כיצד האלגוריתם המתואר במאמר יעזור לנו למצוא רשימת שירים משותפת שתהיה מוסכמת על כמה שיותר מאזינים ע"י השוואה של דימיון בין הדירוגים האישיים של כל מאזין ונשווה בינו לבין הדירוג שנקבל מהאלגוריתם המבוסס על המאמר הקודם [2].

מנושאים הנלמדו בקורס, ומאמר העוסק בשיטות הערכה שונות [7] הסקנו כי הדרך הטובה ביותר להערכת טיב הדירוג החברתי שניצור היא באמצעות שיטה הנקראת $MAP = \text{mean at precision}$. השיטה הזו מחשבת עבור כל משתמש (ולבסוף יוצרת ממוצע על כל המשתמשים) את כמות ההתאמות של הדירוג של כל

שיר בפלייליסט המשותף עם הדירוג שלו אצל המשתמש. כלומר למשל עבור משתמש מסוים שיר x במקום השני ובדירוג החברתי אותו שיר x הגיע למקום שלישי, לכן אין התאמה עבור שיר זה. כך ניתן לבדוק את שביעות הרצון של המשתמש מבלי לערב את רגשותיו וחוויותיו האישיות הסובייקטיביות (למשל חוסר חשק לדרג בצורה אמיתית וכנה או רצון לעזור לחברים לדרג בצורה לא אובייקטיבית את הפלייליסט).

תוכנית פעולה:

<u>משימה</u>	<u>בוצע/לא בוצע</u>	<u>תאריך יעד</u>
מציאת מאגר שירים רלוונטי	בוצע	---
יצירת ממשק משתמש עבור הפלייליסט המשותף ובחירת פלייליסט אישי	בתהליך	25/08/2020
בחירת מדגם מסוים של משתתפים	בוצע	---
מימוש אלגוריתם ראשון	בוצע	---
מימוש אלגוריתם שני	לא בוצע	18/08/2020
הצגת הנתונים הסופיים	לא בוצע	05/09/2020
ניתוח שביעות רצון ודיונים	לא בוצע	08/09/2020

מימוש:

- יצירת רשימת שירים רלוונטית למאגר ראשוני (סגנון זומבה – לקחנו מרשימה פופולארית מוכרת ביוטיוב). הרשימה מורכבת מ-50 שירים. עליו יבצעו המשתמשים את הדירוגים האישיים, אשר אותנו יישמשו לטובת יצירת פלייליסט משותף.
- בניית ממשק אינטרנטי המורכב משני עמודים:
 - עמוד ראשי המכיל את הפלייליסט המשותף המתבסס על תוצאות האלגוריתם של הדירוג החברתי מהדירוגים האישיים. הפלייליסט המשותף ייגזר באופן דינאמי לפי אורך השיעור אשר יוגדר בממשק על ידי המשתמשים.
 - העמוד השני יכיל טופס הכנסת דירוג אישי על הרשימה שהוזכרה בסעיף קודם. המשתמשים לא מחויבים לדרג את כל השירים, האלגוריתם שלנו יודע להתמודד עם דירוגים חלקיים אשר אינם שווים ביניהם.
- העמודים ימומשו באמצעות ממשק flask של פייתון. נרשום את התוצאות של הדירוגים האישיים אל קובץ אקסל וממנו נשאב נתונים עבור האלגוריתמים של הדירוג החברתי.
- הקשר בין שני העמודים יתבצע על ידי אלגוריתם מבחירה חברתית. בחרנו לממש שני אלגוריתמים אשר יפיקו דירוג חברתי של השירים על סמך ההצבעות האישיות ונבחן מי מביניהם ישיג את הדירוג ה"טוב ביותר". נציג כעת את הפסאודו קוד עבור כל אלגוריתם:

Set of songs S

Set of participants P

For each participant $p_j \in P$, a partial ranking \succ_j over $T_j \subseteq S$:

Target playlist G

Output: a partial order $\widehat{\succ}_G$ over S

1. For each $s \in S$ do:
2. $C(s) = \sum_{s_i \in S \setminus S} rv(s, s_i, P)$
3. End for
4. $\widehat{\succ}_G \leftarrow \{\forall (s_i, s_j) \in \binom{S}{2}, s_i \succ_G s_j \text{ iff } c(s_i) > c(s_j)\}$
5. *return* $\widehat{\succ}_G$

Where:

- $rv(s_i, s_j, P) = \text{sign}(\sum_{k \in P} \gamma(s_i, s_j, \succ_k))$
- $\gamma(s_i, s_j, \succ_k) = \begin{cases} 1 & \text{if } s_i \succ s_j \\ -1 & \text{if } s_j \succ s_i \\ 0 & \text{otherwise} \end{cases}$

```

V – number of voters
P – number of preferences
for j in range(0, num_of_voters):
    total_ = 0
    total_2 = 0
    b = (new_b(full_rank[j], num_))
    for i in range(0,V):
        C_ = (new_b(full_rank[i], num_))
        sum1 = ((make_D(B, C, P) + (0.5*(R(B, C) + R(C, B)))))
        sum2 = ((make_D(B, C, P) + (max(R(B, C), R(C, B)))))
        total_ += sum1
        total_2 += sum2
    scores.append(total_ /V)
    scores2.append(total_2/V)

```

where:

- $R(B, C) = \text{sum_Bi}(C) - \text{all_Cij}(C, B)$
- $\text{Sum_Bi}(C)$:


```

total = 0
const = 2
for i in range(1, len(B)+1):
    temp = len(B[i])
    temp2 = math.factorial(temp)
    if temp >= 2 :
        temp3 = math.factorial(temp-const)
        total += (temp2/(temp3*2))

return total

```
- $\text{all_Cij}(C, B)$:


```

total = 0
for j in range (1, len(B)+1):
    for i in range(1, len(C)+1):
        cij = count_of_column_j
        if cij >= 2 :
            total += (math.factorial(cij)/ (math.factorial(cij-2) * 2))

return total

```

- נבצע אבולוציה ברגע שיהיו לנו מספיק נתונים - נחשב לפי שיטת MAP אשר הוזכרה בסקירה הספרותית את הציונים של הדירוגים של תוצאות שני האלגוריתמים ונשווה ביניהם. האלגוריתם בעל הדירוג עם הציון הגבוה ביותר בעצם קלע מדויק יותר אל רצונות המשתמשים ולכן בעל ביצועים טובים יותר.

1. "EduRank: A Collaborative Filtering Approach to Personalization in E-learning", Avi Segal, Ziv Katzir, Ya'akov (Kobi) Gal, Guy Shani, Bracha Shapira, 2014
2. "Computing Distances Between Partial Rankings" Mukul S. Bansal, David Fernandez-Baca ´ Dept. of Computer Science, Iowa State University, Ames, IA 50011, USA, 2008
3. "DJ-MC: A Reinforcement-Learning Agent for Music Playlist Recommendation", Elad Liebman, Maytal Saar-Tsechansky, Peter Stone, 2015
4. "Build your own music recommender by modeling internet radio streams. In Proceedings of the 21st international conference on World Wide Web", N. Aizenberg, Y. Koren, and O. Somekh. pages 1-10. ACM, 2012
5. "Smarter than genius? human evaluation of music recommender systems. In International Symposium on Music Information Retrieval", L. Barrington, R. Oda, and G. Lanckriet. 2009.
6. "The natural language of playlists", B. McFee and G. R. Lanckriet. In ISMIR, pages 537-542, 2011
7. "Gathering and Ranking Photos of Named Entities with High Precision, High Recall, and Diversity". B. Taneva, M. Kacimi and G. Weikum, 2010.
8. Emotional and motivational uses of music in sports and exercise: A questionnaire study among athletes. Petri Laukka and Lina Quick, 2011
9. Social Choice and Multicriterion Decision-Making. Kenneth J. Arrow & Herve Raynaud, 1986.