

מערכת המלצות טקטיות ל-NBA

פרויקט גמר – תוכנית עבודה מפורטת

מידע על הפרויקט

שם הפרויקט: NBA Tactical Recommendations System – חיזוי זמן פסק זמן

אופטימלי

מסלול: אקדמי

שנת לימודים: 2025-2026

1. הקדמה והגדרת הבעיה

פרויקט זה מתמקד בפיתוח מערכת המלצות טקטיות מבוססת למידת מכונה עבור מאמני NBA, המתמקדת בשאלה מרכזית: מתי הזמן האופטימלי לקרוא פסק זמן? החלטה זו נחשבת לאחת ההחלטות הטקטיות החשובות ביותר במהלך משחק ודורשת הבנה עמוקה של מושג ה-מומנטום בכדורסל.

1.1 מניע (Motivation)

מאמני NBA נדרשים לקבל החלטות טקטיות מורכבות בזמן אמת. אחת מהן היא מתי לקרוא פסק זמן כדי לעצור רצף נקודות של היריבה, לארגן מחדש את הקבוצה, או לנצל מצב משחק. כיום, החלטות אלו מתבצעות בעיקר על סמך אינטואיציה וניסיון אישי, ללא תמיכה חישובית או אלגוריתמית. מטרת הפרויקט היא לספק כלי מבוסס מדע שיעזור למקסם את יעילות החלטות פסק הזמן.

1.2 מטרת הפרויקט (Objective)

המטרה המרכזית: לפתח מודל למידת מכונה שמסוגל לחזות את הזמן האופטימלי לפסק זמן, תוך שימוש בהגדרות מחקריות קיימות של מומנטום כמשתנה מרכזי להסבר. שאלת המחקר: האם תזמון הפסק זמן (בהתאם למצב המומנטום) משפיע על יעילותו?

2. אתגרים אלגוריתמיים ומורכבות

2.1 האתגר המרכזי: הגדרת מומנטום

מומנטום הוא מושג מורכב ורב-ממדי, ללא הגדרה מוסכמת בספרות הספורטיבית. האתגר האלגוריתמי המרכזי בפרויקט הוא:

- **הגדרה כמותית:** כיצד להמיר מושג מופשט ("מומנטום") למדד מדיד וחישובי?
- **ריבוי משתנים:** מומנטום מושפע מפרמטרים רבים – רצפי נקודות, הגנה, עיפות שחקנים, עבירות, קהל וכו'.
- **תלות בזמן:** מומנטום משתנה דינמית במהלך המשחק – מה שהיה "מומנטום חיובי" לפני שתי דקות עלול להפוך לשלילי.
- **משתנים נסתרים:** ישנם גורמים שלא ניתן למדוד ישירות (מורל הקבוצה, לחץ פסיכולוגי).
- **אי-ליניאריות:** הקשר בין משתנים למומנטום אינו ליניארי – לדוגמה, רצף 0-6 בדקה האחרונה משפיע אחרת מאשר רצף 0-6 בתחילת המשחק.

2.2 גישתנו למורכבות

הפרויקט מתמודד עם מורכבות זו על ידי שילוב מחקר קיים עם עבודה אלגוריתמית מקורית:

- **שימוש בהגדרות מומנטום מבוססות מחקר:** ניתוח מאמרים אקדמיים והגדרות קיימות של מומנטום כנקודת פתיחה.
 - **מודלים סטטיסטיים מתקדמים:** שימוש ב-Machine Learning (Random Forest, Gradient Boosting) לזיהוי דפוסים מורכבים.
 - **פיתוח משתנים נגזרים (Feature Engineering):** יצירת משתנים מתוחכמים כגון "רצף נקודות משוקלל בזמן", "שינוי קצב משחק".
 - **גישת חיזוי רטרואספקטיבית:** למידה מנתונים היסטוריים – בחינת יעילות פסקי זמן שכבר נלקחו.
-

3. רכיב למידת המכונה (Machine Learning)

למידת מכונה היא הליבה האלגוריתמית של הפרויקט. מדובר בתהליך מחקרי מלא הכולל:

3.1 תהליך הלמידה

- **איסוף ועיבוד נתונים:**

- חילוץ נתוני Play-by-Play מ-NBA API
- זיהוי אירועי פסק זמן והקשרם בזמן
- ניקוי נתונים וחסרים, טיפול בערכים חריגים

- **פיתוח משתנים (Feature Engineering):**

- חישוב פערי נקודות לאורך זמן
- זיהוי רצפי נקודות (Consecutive Scoring Runs)
- חישוב ציון מומנטום על בסיס מחקר
- יצירת משתנים נגזרים: קצב משחק, עייפות מצטברת, עבירות

- **בחירת מודלים ואופטימיזציה:**

- בדיקה של מספר אלגוריתמים: Logistic Regression, Random Forest, XGBoost
- חיפוש גריד/רנדומלי (Grid/Random Search) להתאמת היפר-פרמטרים
- Cross-validation למניעת Overfitting

- **הערכת ביצועים:**

- מדדי Precision, Recall, F1-Score
- ROC Curve ו-AUC

3.2 חידוש אלגוריתמי

תרומת המחקר של הפרויקט היא בחינת הקשר בין מומנטום לזמן קריאת פסק זמן, תחום שלא נחקר באופן שיטתי במדעי הספורט:

- האם פסק זמן ברגע "הנכון" (כאשר מומנטום נגדנו) עוצר באמת את רצף הנקודות של היריבה?
- האם קיימת "חלון זמן אופטימלי" לקריאת פסק זמן?
- האם פסק זמן מוקדם או מאוחר מדי פחות יעיל?
- מהו ההבדל בין סוגי מומנטום שונים (נקודות, הגנה, קהל) בהקשר פסקי זמן?

4. רכיב בזמן אמת: צ'אטבוט למאמנים

כחלק מדרישות הפרויקט לפיתוח בזמן אמת, יפותח צ'אטבוט אינטראקטיבי שמאפשר למאמנים (או לעוזרי מאמן) לקבל המלצות מיידיות במהלך המשחק.

4.1 ארכיטקטורה

המערכת בנויה בשני שלבים:

1. **אימון (Offline):** המודל מאמן על נתונים היסטוריים ולומד דפוסי אופטימליות.
2. **חיזוי בזמן אמת (Real-Time):** המאמן מזין פרמטרים נוכחיים בצ'אטבוט, והמודל מחזיר המלצה תוך שניות.

4.2 ממשק משתמש

הצ'אטבוט כולל שאלות מוכנות מראש, כאשר המשתמש מזין רק את הערכים:

- רבע נוכחי וזמן שנותר

- פער נקודות
 - רצף נקודות היריבה: X-X ב-Y דקות
 - רצף נקודות הקבוצה שלנו
 - עבירות של שחקנים מרכזיים
 - פסקי זמן שנותרו
 - הערות נוספות (טקסט חופשי – אופציונלי)
- יתרון הגישה: חיזוי אמיתי בזמן אמת – המודל פועל על נתונים נוכחיים. הכנסת הנתונים ידנית אינה פוגעת בלגיטימיות, והגישה פרקטית ונוחה לשימוש.
-

5. מקורות נתונים ותשתית טכנית

5.1 מקורות ראשיים

- **NBA API** (חבילת `python nba_api`):
 - נתוני Play-by-Play מפורטים לכל משחק
 - זיהוי אירועי פסק זמן (`EVENTMSGTYPE`)
 - ניקוד, פער נקודות, זמן משחק
 - נתונים היסטוריים לאורך עשורים
- **Basketball-Reference** (תוספתי):
 - סטטיסטיקות שחקנים מתקדמות
 - מדדים מתקדמים: PER, Win Shares וכו'

5.2 סביבת פיתוח

- +Python 3.7
 - nba_api, pandas, numpy, scikit-learn
 - matplotlib, seaborn
 - Jupyter Notebooks
 - Flask/Streamlit
-

6. תוכנית עבודה מפורטת

שלב 1 – הקמה וסביבת עבודה (שבוע 1):

- התקנת חבילות Python נדרשות
- בדיקת חיבור ל-NBA API והורדת נתוני דוגמה
- הקמת מבנה פרויקט מאורגן
- קריאה ראשונית של מאמרי מומנטום

שלב 2 – מחקר והגדרת הבעיה (שבועות 2-3):

- מציאת 5–10 מאמרים רלוונטיים
- הגדרת מומנטום – איזו הגדרה נשתמש
- הגדרת "פסק זמן מוצלח" – קריטריונים להצלחה
- זיהוי פערי ידע במחקר הקיים
- כתיבת סקירת ספרות

שלב 3 – איסוף ועיבוד נתונים (שבועות 4-6):

- הורדת נתוני Play-by-Play ממספר עונות
- זיהוי כל אירועי פסק זמן
- חישוב פערי נקודות לאורך זמן
- זיהוי רצפים (Runs) – אלגוריתם לזיהוי רצפי ניקוד
- ניקוי נתונים וטיפול בחסרים
- יצירת בסיס נתונים אחיד ונקי

שלב 4 – Feature Engineering (שבועות 7-8):

- חישוב ציון מומנטום לפי מחקר
- יצירת משתנים נגזרים: קצב משחק, עייפות, מומנטום משוקלל
- ניתוח מתאמים בין משתנים
- בחירת משתנים חשובים
- יצירת ויזואליזציות להבנת הנתונים

שלב 5 – בניית מודלי ML (שבועות 9-11):

- חלוקת הנתונים ל-Train/Validation/Test
- בניית מודל בסיסי (Logistic Regression)
- בניית מודלים מתקדמים (Random Forest, XGBoost)
- חיפוש היפר-פרמטרים
- Cross-validation והערכת ביצועים
- ניתוח שגיאות

שלב 6 – פיתוח צ'אטבוט (שבועות 12-14):

- עיצוב ממשק משתמש
- בניית שאלות מוכנות מראש
- אינטגרציה עם המודל המאומן
- בדיקה ואופטימיזציה של זמני תגובה
- הוספת יכולות שיחה נוספות (הסברים, היסטוריה)

שלב 7 – בדיקות ואימות (שבועות 15-16):

- בדיקות יחידה (Unit Testing)
- בדיקות אינטגרציה
- אימות תוצאות מול משחקים ידועים
- בדיקות שימושיות
- תיקוני באגים ושיפורים

שלב 8 – תיעוד והשלמה (שבועות 17-18):

- כתיבת דוח סופי
- הכ

נת מצגת ביצוע

- הכנת פוסטר ליום ציבורי
 - הכנת הדגמות
 - יצירת סרטון הסבר (אופציונלי)
-

7. מדדי הצלחה ובקרת איכות

7.1 מדדים כמותיים

- Precision: יעד $\leq 75\%$
- Recall: יעד $\leq 70\%$
- F1-Score: יעד ≤ 0.7
- AUC-ROC: יעד ≤ 0.75
- זמן תגובת הצ'אטבוט: יעד > 3 שניות

7.2 מדדים איכותיים

- המודל מסוגל להסביר את החלטותיו (Interpretability)
- הצ'אטבוט ידידותי למשתמש
- תוצאות עקביות עם ידע מחקרי קיים
- מערכת מספקת ערך ממשי למאמנים

8. סיכום ודרישות הקורס

פרויקט זה משלב את כל רכיבי הפרויקט האקדמי:

- הגדרת הבעיה: הבנה עמוקה של השאלה האופטימלית לפסק זמן בהקשר מומנטום
- אלגוריתמים: פתרון מקורי באמצעות ML, עם טיפול מורכב בהגדרת מומנטום
- עיצוב: עיצוב מערכת רחב הכולל איסוף נתונים, מודל ML וממשק משתמש
- תכנות: תוכנה רב-שכבתית מורכבת

- בדיקת תוצאות: הערכה כמותית ואיכותית
- למידה עצמית: חקירה מעמיקה של Data Science, ML, NBA API ומחקרי ספורט
- עבודה שיטתית: תוכנית עבודה מפורטת עם בקרת איכות בכל שלב

שתי שאלות מרכזיות שהפרויקט יטפל בהן:

1. כיצד להגדיר מומנטום באופן כמותי ומדויק?
2. מה מגדיר פסק זמן "מוצלח"? – קריטריון זה יקבע כיצד נאמן ונעריך את המודל

הערה: תוכנית עבודה זו היא מסמך חי, שתעודכן בהתאם להתקדמות הפרויקט, ממצאי מחקר והנחיית היועץ.
