

# 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 בשאלת מרכזית: **מתי הזמן האופטימלי לקרוא פסק זמן?** החלטה זו נחשבת לאחר החלטות הטקטיות החשובות ביותר במהלך משחק ודורשת הבנה עמוקה של מושג **המוננטום** בכדורסל.

### 1.1 מניע (Motivation)

מאמני NBA נדרשים לקבל החלטות טקטיות מורכבות בזמן אמיתי. אחת מהן היא מתי לקרוא פסק זמן כדי לעזור נזודות של היריבה, לארגן מחדש את הקבוצה, או לנצל מצב משחק. כיום, החלטות אלו מתבצעות בעיקר על סמך אינטואיציה וניסיון אישי, ללא תמייכה חישובית או אלגוריתמית. מטרת הפרויקט היא **לספק כלי מבוסס מדע** שיעזר 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 בהגדלות מחקריות קיימות של מומנטום כמשתנה מרכזי להסביר. שאלת המחקר: האם בזמן פסק זמן (בהתאם למצב המומנטום) משפייע על ייעולו?

## 2. אתגרים אלגוריתמיים ומורכבות

### 2.1 האתגר המרכזי: הגדרת מומנטום

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

- **הגדרה כמותית:** כיצד להמיר מושג מופשט ("מומנטום") למדד מדייד וחישובי?
- **ריבוי משתנים:** מומנטוםמושפע מפרמטרים רבים – רצפי נקודות, הגנה, עייפות שחקנים, עבירות, קהל וכו'.
- **תלות בזמן:** מומנטום משתנה דינמית במהלך המשחק – מה שהוא "מומנטום חיובי" לפני שתי דקוטות עלול להפוך לשילייל.
- **משתנים נסתרים:** ישם גורמים שלא ניתן למדוד ישירות (מורל הקבוצה, לחץ פסיקולוגי).
- **אידיאליות:** הקשר בין משתנים למומנטום אינו ליניארי – לדוגמה, רצף 0-6 בדקה الأخيرة משפייע אחרת מאשר רצף 0-6 בתחילת המשחק.

### 2.2 גישתנו למורכבות

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

- **שימוש בהגדרות מומנטום מבוססות מחקר:** ניתוח מאמרים אקדמיים והגדרות קיימות של מומנטום כנקודות פתיחה.
- **מודלים סטטיסטיים מתקדמים:** שימוש ב-Random Forest, Gradient Boosting
- **פיתוח משתנים נגזרים (Feature Engineering):** יצירת משתנים מתחכמים כגון "רצף נקודות משוקלל בזמן", "שינוי קצב משחק".
- **גישה חיזוי רטרוספקטיבית:** למדיה מנתונים היסטוריים – בחינת ייעילות פסק זמן שכבר נלקחו.

## **3. רכיב למידת המquina (Machine Learning)**

למידת מכונה היא הליבה האלגוריתמית של הפרויקט. מדובר בתהיליך מחקרי מלא הכלול:

### **3.1 תהיליך הלמידה**

- איסוף ועיבוד נתונים:**

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

- פיתוח משתנים (Feature Engineering):**

- חישוב פערן נקודות לאורך זמן
- זיהוי רצפי נקודות (Consecutive Scoring Runs)
- חישוב ציון מוננטום על בסיס מחקר
- ייצור 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) להסתאמת היפר-פרמטרים
- Overfitting Cross-validation

- הערכת ביצועים:**

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

- ניתוח שגיאות – מתי המודל נכשל ולמה

## 3.2 חידוש אלגוריתמי

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

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

## 4. רכיב בזמן אמת: צ'אבטות למאמנים

חלק מדרישות הפרויקט לפיתוח בזמן אמת, יפותח צ'אבטות אינטראקטיבי שמאפשר למאמנים (או לעוזרי מאمن) לקבל המלצות מיידיות במהלך המשחק.

### 4.1 ארכיטקטורה

המערכת בנויה בשני שלבים:

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

### 4.2 ממשק משתמש

הצ'אבטות כולן שאלות מוכנות מראש, כאשר המשתמש מזין רק את הערכים:

- רביע נוכחי וזמן שנוצר

- פער נקודות
- רצף נקודות היריבה: X-X ב-Y דקota
- רצף נקודות הקבוצה שלנו
- עבירות של שחקנים מרכזים
- פסק זמן שנותרו
- הערות נוספת (טקסט חופשי – אופציונלי)

יתרון הגישה: חיזוי אמיתי בזמן אמת – המודל פועל על נתונים הנוכחיים. הכנסת הנתונים ידנית אינה פוגעת בלוגיטיות, והגישה פרקטית ונוחה לשימוש.

---

## 5. מקורות נתונים ותשתיות טכנית

### 5.1 מקורות ראשיים

- API NBA (חברת nba\_api):

  - נתונים Play-by-Play מפורטים לכל משחק

  - זיהוי אירוע פסק זמן (EVENTMSGTYPE)

  - ניקוד, פער נקודות, זמן משחק

  - נתונים היסטוריים לאורך עשורים

- Basketball-Reference (תוספות):

  - סטטיסטיות 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 נדרשת
- בדיקת חיבור ל-API NBA והורדת נתונים 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" – קритרונים להצלחה
- זיהוי פער ידע במחקר הנוכחי
- כתיבת סקירת ספרות

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

- הורדת נתונים Play-by-Play מספר עונות

- זיהוי כל אירוע פסק זמן

- חישוב פער נקודות לאורך זמן

- זיהוי רצפים (Runs) – אלגוריתם לזיהוי רצפי ניקוד

- ניקוי נתונים וטיפול בחסרים

- ייצור בסיס נתונים אחיד ונקי

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

- חישוב ציון מומנטום לפי מחקר

- ייצור משתנים נגזרים: קצב משחק, עייפות, מומנטום משוקל

- ניתוח מתאימים בין משתנים

- בחירת משתנים חשובים

- ייצור ויזואלייזציות 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 זמן תגובה
- הוספת יכולות שיחה נוספת (הסברים, היסטוריה)

#### **שלב 7 – בדיקות ואיימות (שבועות 15-16):**

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

#### **שלב 8 – תיעוד והשלמה (שבועות 17-18):**

- כתיבת דוח סופי
- הц

#### **נת מצגת ביצוע**

- הכנת פווטר ליום ציבורי
- הכנת הדגמות
- ייצור סרטון הסבר (אופציונלי)

## 7. מדדי הצלחה וביקורת איות

### 7.1 מדדים כמותיים

•  $\text{Precision} \leq 75\%$ : יעד

•  $\text{Recall} \leq 70\%$ : יעד

•  $\text{F1-Score} \leq 0.7$ : יעד

•  $\text{AUC-ROC} \leq 0.75$ : יעד

•  $\text{זמן תגובה הצ'אטיבוט}: \text{יעד} > 3 \text{ שניות}$

### 7.2 מדדים איותיים

• המודל מסוגל להסביר את החלטותיו (Interpretability)

• הצ'אטיבוט ידידותי למשתמש

• תוצאות עקביות עם ידע מחקרי קיימ

• מערכת מספקת ערך ממשי לאמנים

---

## 8. סיכום ודרישות הקורס

פרויקט זה משלב את כל רכיבי הפרויקט האקדמי:

- הגדרת הבעיה: הבנה عمוקה של השאלה האופטימלית לפסק זמן בהקשר מומנטום
- אלגוריתמים: פתרון מקורי באמצעות LM, עם טיפול מורכב בהגדרת מומנטום
- עיצוב: עיצוב מערכת רחב הכליל איסוף נתונים, מודל LM וממשק משתמש
- תכנונות: תוכנה רב-שכבותית מורכבת

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

**שתי שאלות מרכזיות שהפרויקט יטפל בהן:**

1. כיצד להגדיר מומנטום באופן איכותי ומדויק?
2. מה מגדיר פסק זמן "מושלח"? – קритריון זה יקבע כיצד נאמן ונעריך את המודל

הערה: תוכנית עבודה זו היא מסמר ח', שתעדכן בהתאם להתקדמות הפרויקט, ממצאי מחקר והנחיית היועץ.

---