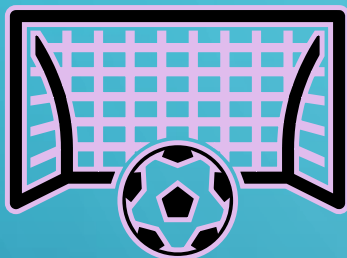


סדנת הכנה לפרויקט בהנדסת מערכות מידע

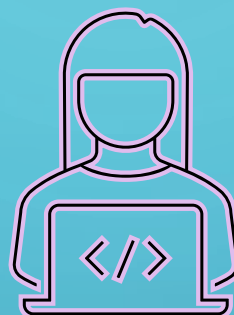
איילת מויאל, נטליה קטייב, ליאל בנימין, יונתן גרינשפן

מוטיבציה



המטרה

כיום אנו מסוגלים לחזות תוצאות של משחקי ספורט רבים, וטכנולוגיה זו הופכת מדויקת יותר ומשתפרת.



מחקר לפתרון בעיות סיווג וחיזוי

שימוש במודלים שונים על מנת לחזות ולפתור בעיות סיווג



בניית מערכת לניהול ליגות כדורגל

כדורגל הוא הספורט הפופולרי בעולם, והענף של חיזוי תוצאות כדורגל הולך וגדל במרוצת השנים.

שלבים בתהליך העבודה

סקירת תחום וספרות

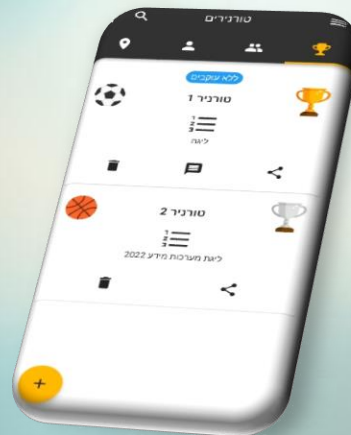
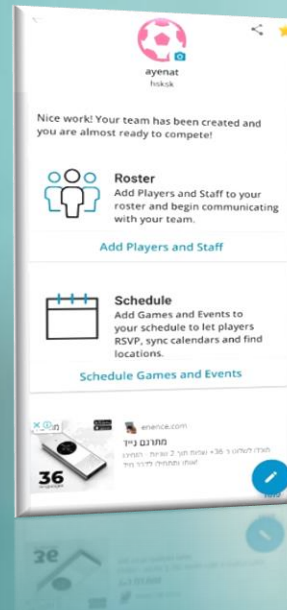
אפיון ותיכון המערכת

מימוש המערכת

ניתוח נתונים וחיזוי

סקירת התחום וסקירה מעשית

1. ביצענו סקירה מעשית על תחום ניהול ליגות
2. ביצענו סקירה על מערכות קיימות לחיזוי תוצאות משחקים.



דוגמא להשוואה בין המערכות הבאות לניהול ליגות הקיימות שאותם סקרנו:

קריטריון להשוואה/ פיצ'ר מרכזי	Winner	Monclubsportif	Sports engine
חישוב טבלת ליגה	• קיים.	• קיים.	• קיים.
התראות על משחקים ואירועים	<ul style="list-style-type: none"> • אין אפשרות לשלוח הודעות מתוזמנות לעוקבי (ואוהדי) הליגה. • מאפשר שליחת הודעה ידנית לכל עוקבי (ואוהדי) הליגה שמקפיצה להם הודעת push. 	<ul style="list-style-type: none"> • מאפשר הגדרת התראות אוטומטיות לקבוצות ולעוקבים (ואוהדים) בהתאם לבחירת המשתמש (ניתן לקבל התראה באמצעות אימייל/הודעת push/שניהם ביחד). 	<ul style="list-style-type: none"> • מאפשר הגדרת התראות אוטומטיות לאוהדים על משחקים אך אין התראות לשופטים.
שיבוץ משחקים	<ul style="list-style-type: none"> • שיבוץ המשחקים נעשה באופן אוטומטי ורנדומלי או באופן ידני ע"י רשימה. (במידה והמשתמש בחר שיבוץ באופן אוטומטי רנדומלי המערכת משבצת משחקים בין כל שתי קבוצות בהתאם למספר המפגשים שאחראי הליגה בוחר) 	<ul style="list-style-type: none"> • שיבוץ המשחקים נעשה באופן ידני לבחירת המשתמש. עבור כל קבוצה מוצגות האפשרויות בלוח שנה. 	<ul style="list-style-type: none"> • שיבוץ המשחקים נעשה באופן ידני לבחירת המשתמש. עבור כל קבוצה מוצגות האפשרויות בלוח שנה.
בעלי תפקידים אפשריים במערכת	• שחקן.	<ul style="list-style-type: none"> • שחקן. • שחקן מחליף. • מאמן. 	<ul style="list-style-type: none"> • שחקן. • עובד צוות.

דוגמא למערכות קיימות לחיזוי תוצאות משחקים:

קריטריון להשוואה	Forebet	SoccerVista
פלטפורמות בהן נתמך	<ul style="list-style-type: none"> • Web. • Mobile. 	<ul style="list-style-type: none"> • Web.
מדינות בהן ניתן להשתמש	<ul style="list-style-type: none"> • עולמי. • מעל 400 ליגות. 	<ul style="list-style-type: none"> • עולמי.
שנת הקמת האתר	• 2009	• 1999
עלות שימוש	• חינמי	• חינמי
נתונים נגזרים	<ul style="list-style-type: none"> • האם יהיו מעל או מתחת ל2.5. • שערים במשחק. • האם שתי הקבוצות יבקיעו במשחק העתידי. • האם הקבוצה תנצח במשחק או שיהיה תיקו. 	<ul style="list-style-type: none"> • האם יהיו מעל או מתחת ל2.5. • שערים במשחק. • האם הקבוצה תנצח במשחק או שיהיה תיקו.
פיצ'רים נחזים מרכזיים	<ul style="list-style-type: none"> • חיזוי של תוצאת משחק סופית. • חיזוי של תוצאת מחצית המשחק. • למי מהשחקנים המשתתפים במשחק יש סיכוי להבקיע ומה הסיכוי. 	<ul style="list-style-type: none"> • חיזוי תוצאת משחק סופית.
האינפורמציה שמוצגת עבור כל משחק	<ul style="list-style-type: none"> • המלצות למשתמש בזמן אמת. • תנאי מזג האוויר במיקום בו מתקיים המשחק. • פירוט אודות אילו שחקנים משתתפים, אילו מושעים ואילו שחקנים פצועים. • הצגת יחסי ההימורים. 	<ul style="list-style-type: none"> • הצגת יחסי ההימורים.

סקירת ספרות

חקרנו על שיטות ואלגוריתמים שונים לפתרון בעיות
Classification.

מודלים ואלגוריתמים:

- ANN
- KNN
- Random Forest
- Decision Tree

ביצענו השוואה בין האלגוריתמים, עבור כל אלגוריתם
בדקנו מה הבעיה שהוא בא לפתור, איך הוא פותר אותה,
מה רמת הדיוק, גודל סט הנתונים וסוג הנתונים.



אפיון ותיכון המערכת

משימה זו הניחה את התשתית והיותה את הבסיס לפרויקט שלנו.
בניית המודלים הבאים:

- Class models
- Sequence diagrams
- Use cases
- Architecture diagram
- Acceptance tests



מעבר מדרישות פונקציונליות לתרחישי שימוש

אורח:

- 2.1. משתמש המבקר במערכת מוגדר כאורח.
- 2.2. אורח יכול לבצע **רישום למערכת** ע"י הזנת פרטים מזהים וקביעת סיסמה בסיום תהליך רישום מוצלח האורח הופך לאוהד במערכת. אבל, כדי להיות מזוהה כמנוי עליו להיכנס (login) למערכת.
- 2.3. **כניסה (login)** למערכת באמצעות פרטים מזהים וסיסמה. בסיום תהליך כניסה מוצלח המשתמש מזוהה כמנוי.



- **תרחיש אלטרנטיבי 1:**
חלופה מצעד 4 של תרחיש הצלחה ראשי: פרטי ההתחברות שהוזנו אינם תקינים.
 - 4.1. המערכת מציגה הודעה כי הפרטים שהוזנו אינם תקינים.
 - 4.2. האורח ממלא את השדות המתאימים מחדש.
 - 4.3. המערכת חוזרת לשלב 4 בתרחיש ההצלחה הראשי.
- **תרחיש אלטרנטיבי 2:**
חלופה מצעד 3 של תרחיש הצלחה ראשי ומצעד 4.2 של תרחיש אלטרנטיבי 1:
המשתמש מתחרט ואינו מעוניין להתחבר למערכת.
 - 3.1. המשתמש לוחץ על כפתור "ביטול".
 - 3.2. המערכת מחזירה את המשתמש בתור משתמש אורח למסך הראשי.



4. **התחברות למערכת – כל משתמש**
- **שחקן ראשי – משתמש רשום למערכת**
- **בעלי עניין ואינטרסים:**
 - ✦ **משתמש רשום** – משתמש המעוניין לבצע תהליך התחברות למערכת בצורה קלה, מהירה ונוחה.
- **תנאי קדם:**
 - ✦ המשתמש רשום למערכת.
- **תנאי סיום:**
 - ✦ התחברות המשתמש בוצעה בהצלחה.
 - ✦ תיעוד ההתחברות נרשם לקובץ הלוג.
- **תרחיש הצלחה ראשי:**
 1. המשתמש בוחר באפשרות של כניסה למערכת.
 2. המערכת מציגה לאורח שדות למילוי של שם משתמש וסיסמה שלו.
 3. המשתמש ממלא את פרטיו בהתאם.
 4. המערכת מבצעת וידוא ואימות של שם המשתמש וסיסמה.
 5. המערכת רושמת תיעוד של ההתחברות לקובץ הלוג.
 6. המערכת מציגה למשתמש הודעה כי ההתחברות בוצעה בהצלחה.
 7. המערכת משנה את סטטוס המשתמש למחובר.
 8. המערכת מעבירה את המשתמש חזרה למסך הראשי.

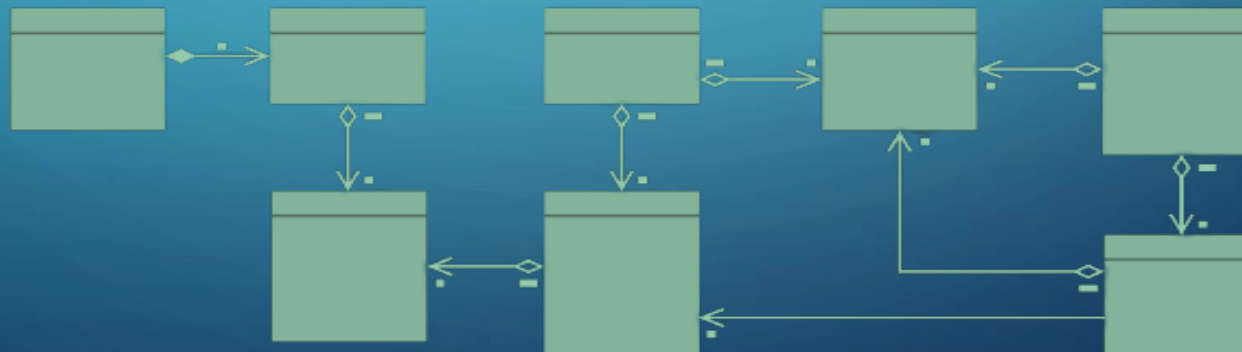
WHITE DIAGRAM & CLASS DIAGRAM

Class Diagram

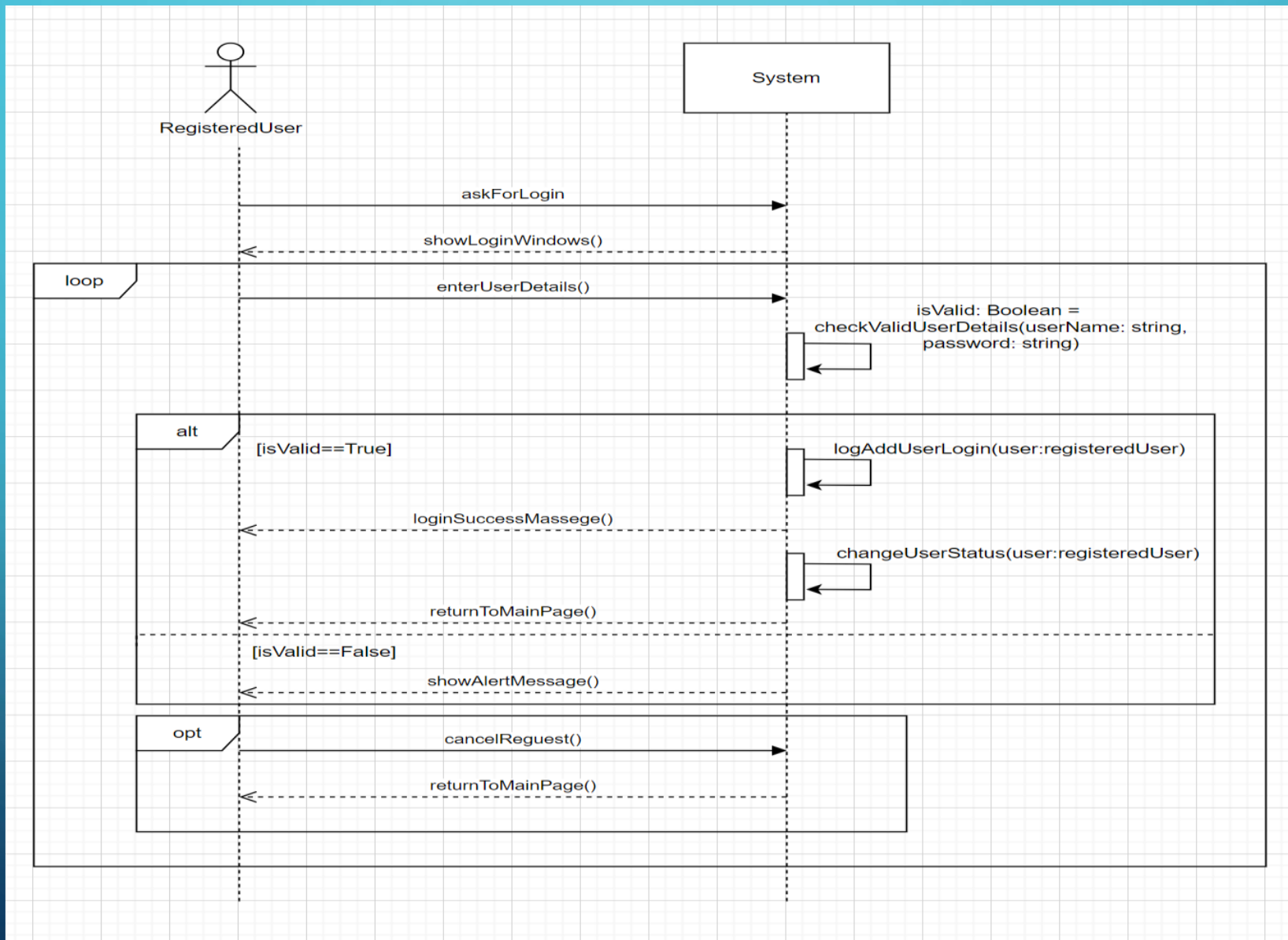
- ★ ★ שיקוף מלא של מידול המערכת
- ★ ★ מעשירה את ה White Diagram
- ★ ★ כולל גם קשרים מכוונים, מטודות, תכונות של ישויות וממשקים

White Diagram

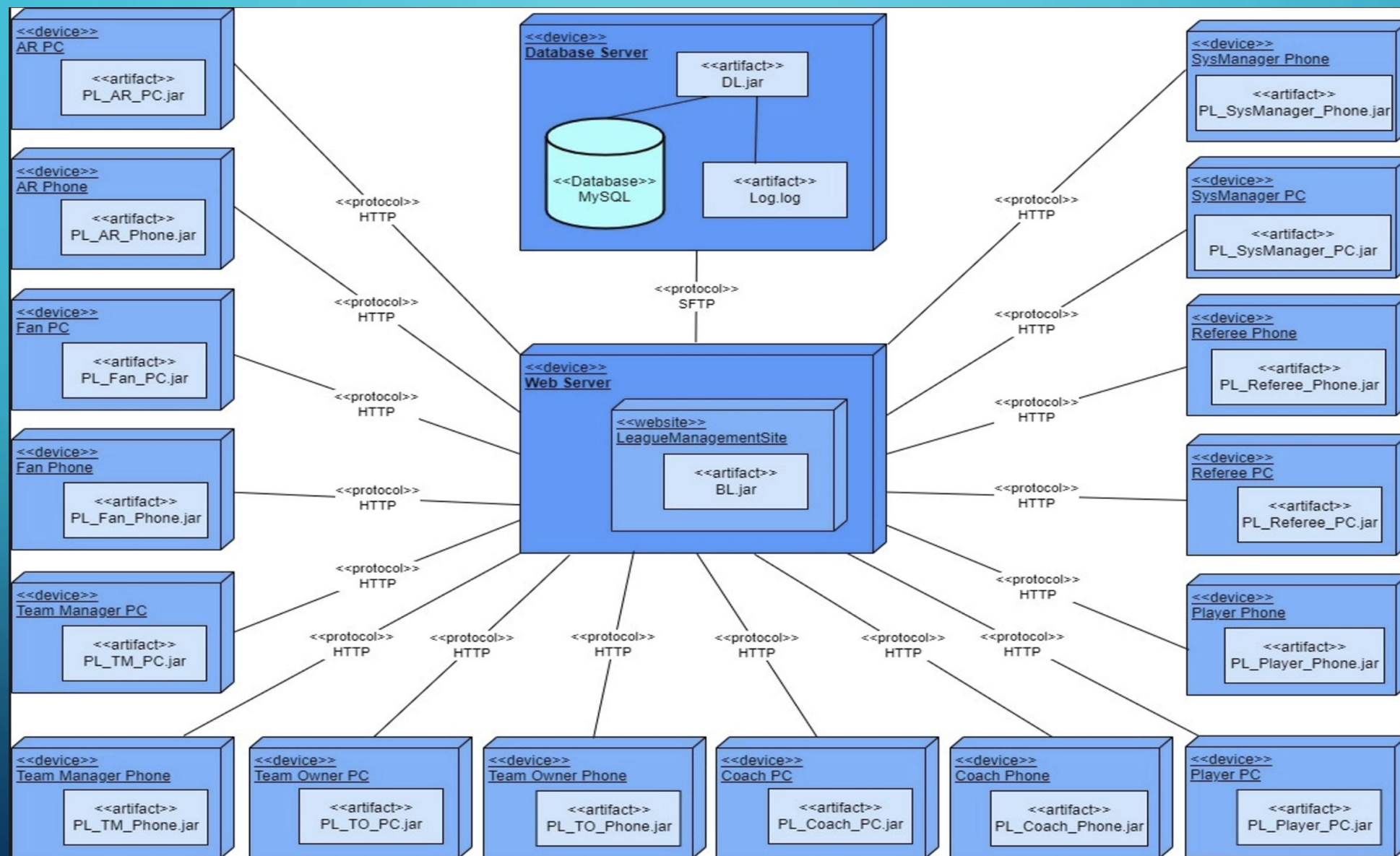
- ★ ★ מבט סטטי על המערכת
- ★ ★ תרשים ראשוני
- ★ ★ מייצג ישויות ואת הקשרים ביניהן



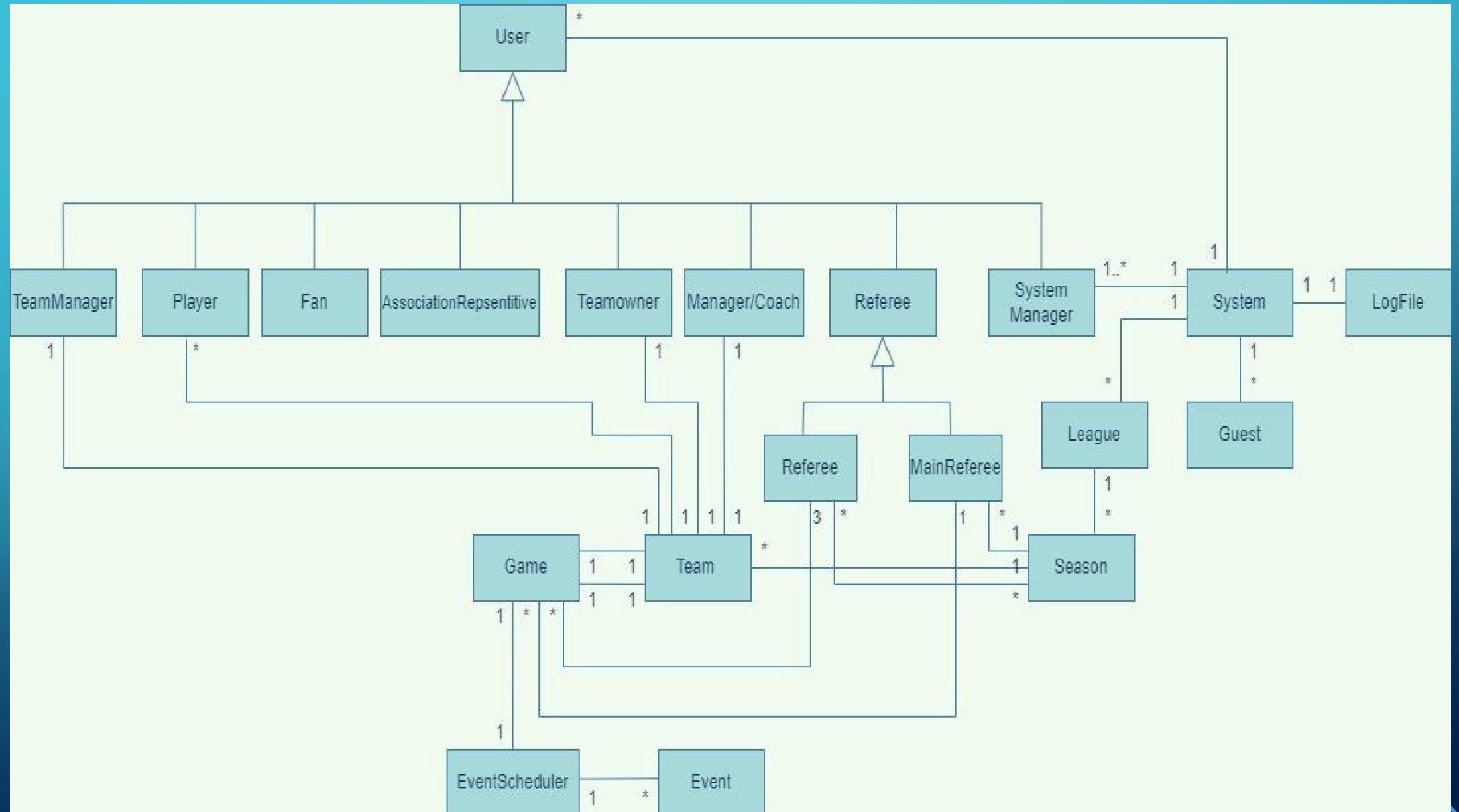
תרשים רצף של תרחיש שימוש התחברות למערכת



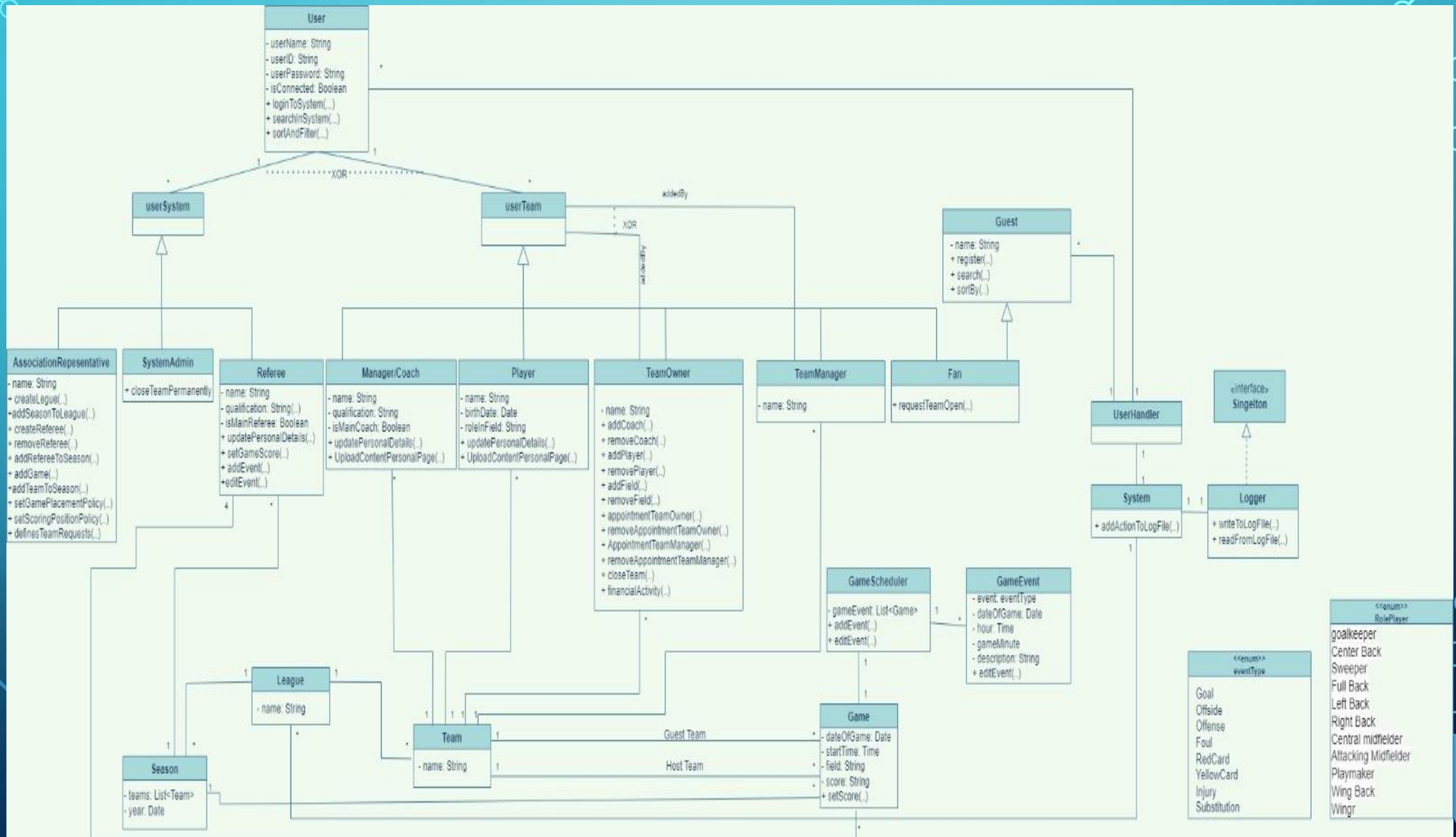
ארכיטקטורת המערכת



White Diagram



Class Diagram



מודל 3 השכבות



שכבת הגישה
(Data Base)



שכבת התחום
(Domain)

שכבת השירות
(Service)



GitHub

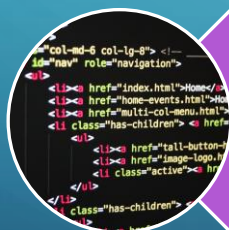
אופן המימוש



מימוש על בסיס אפיון הדרישות



עבודה מול GitHub



תהליך עבודה קפדני – סדר
בקוד, דוקומנטציה, שמירה על
לכידות גבוהה וצימוד נמוך.

בדיקות

Unit tests

Integration
tests

Acceptance
tests



מטריצת עקיבות

מספר בדיקה	תיאור הדרישה	מחלקה אחראית	מס' הדרישה כפי שמופיע במסמך הדרישות	מספר ה-UC
1.1	רישום למערכת	אורח	2.2	1
1.2				
2.1	חיפוש במערכת	משתמש שרשום למערכת	2.5	2
2.2				
3.1	מיון וסינון של תוצאות חיפוש	אורח ומשתמש שרשום למערכת	2.5	3
3.2				
4.1	התחברות למערכת	משתמש שרשום למערכת	2.3	4
4.2				
4.3				

בדיקות קבלה – UAT4

4.1 חיובי:

צעדים:

- ❖ המשתמש בוחר באפשרות של "כניסה למערכת".
- ❖ המשתמש מזין את הפרטים: id=natalia, password=12345

תוצאות:

- ❖ המערכת מוסיפה את תיעוד ההתחברות לקובץ הלוג.
- ❖ המערכת מציגה הודעה למשתמש כי הכניסה למערכת בוצעה בהצלחה.
- ❖ המשתמש מחובר למערכת ויכול להשתמש בה לפי הרשאותיו.

4.2 שלילי:

צעדים:

- ❖ המשתמש בוחר באופציה של "כניסה למערכת".
- ❖ המשתמש מזין את הפרטים: id=notRegistered, password=12345

תוצאות:

המערכת מציגה למשתמש הודעת שגיאה כי הפרטים אינם מזוהים במערכת.

4.3 שלילי:

צעדים:

- ❖ המשתמש בוחר באופציה של "כניסה למערכת".
- ❖ המשתמש מתחרט ואינו רוצה להתחבר ולוחץ על כפתור "ביטול".

תוצאות:

המערכת מחזירה את המשתמש חזרה אל מסך הראשי בתור אורח.

ניתוח נתונים ומודלי חיזוי



ניקוי נתונים

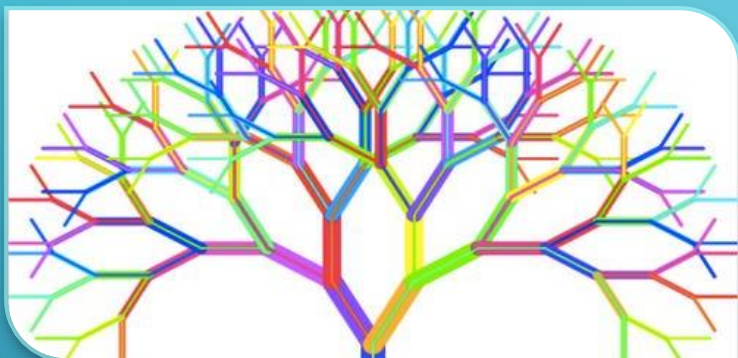
הורדנו רשומות משוכפלות , ערבנו
את הטבלאות ובדקנו עם קיים לנו
מידע חסר



DATASETS

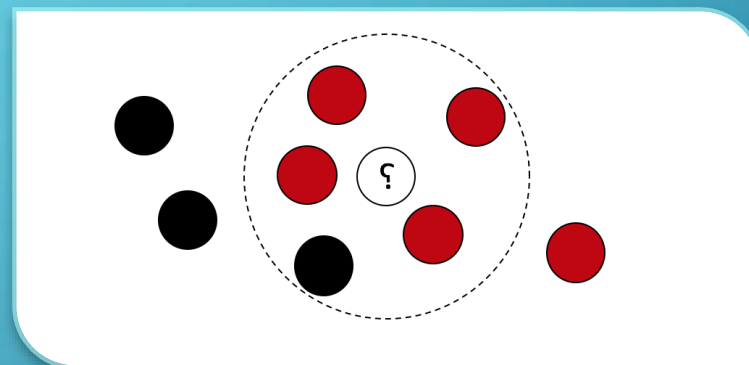
בחרנו 10 מתוך הרשימה

האלגוריתמים שלנו



RANDOM FOREST

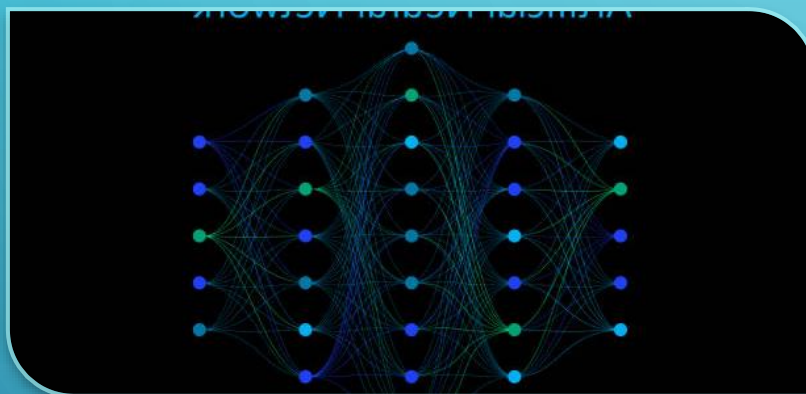
אלגוריתם לסיווג שיוצר מספר עצי החלטה בזמן האימון.



KNN

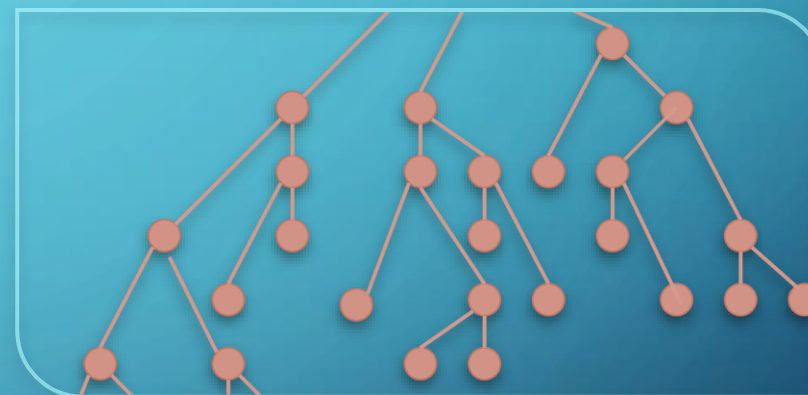
אלגוריתם חסר פרמטרים לסיווג הקלט תלוי ב-K השכנים הקרובים ביותר.

האלגוריתמים שלנו



ARTIFICIAL NEURAL
NETWORK

מודל מתמטי שפותח בהשראת
תהליכים מוחיים, משמש ללמידת
מכונה.



DECISION TREE

מודל חיזוי סטטיסטי המספק מיפוי
בין התצפיות לערכים המתאימים.

היפר - פרמטרים

- עבור אלגוריתם ה KNN בדקנו את הפרמטרים:

- סוג אלגוריתם, מספר שכנים, משקולות וגודל העלה.

- עבור אלגוריתם Random forest בדקנו את הפרמטרים:

- סוג הפיצ'ר ואת הסוג משקולת.

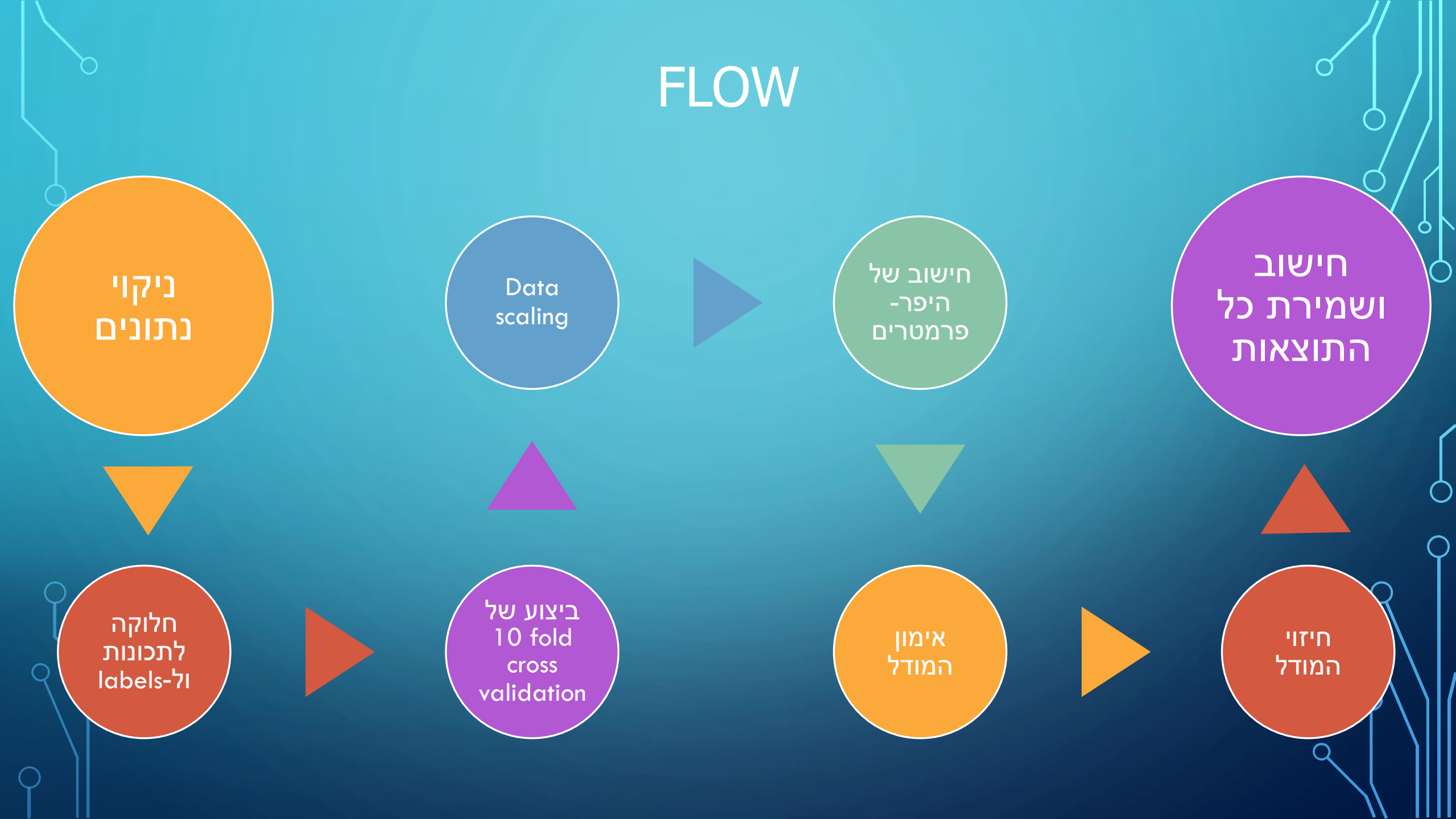
- עבור Decision tree בדקנו את הפרמטרים:

- קריטריון, התפצלות, סוג הפיצ'ר ואת סוג המשקולת.

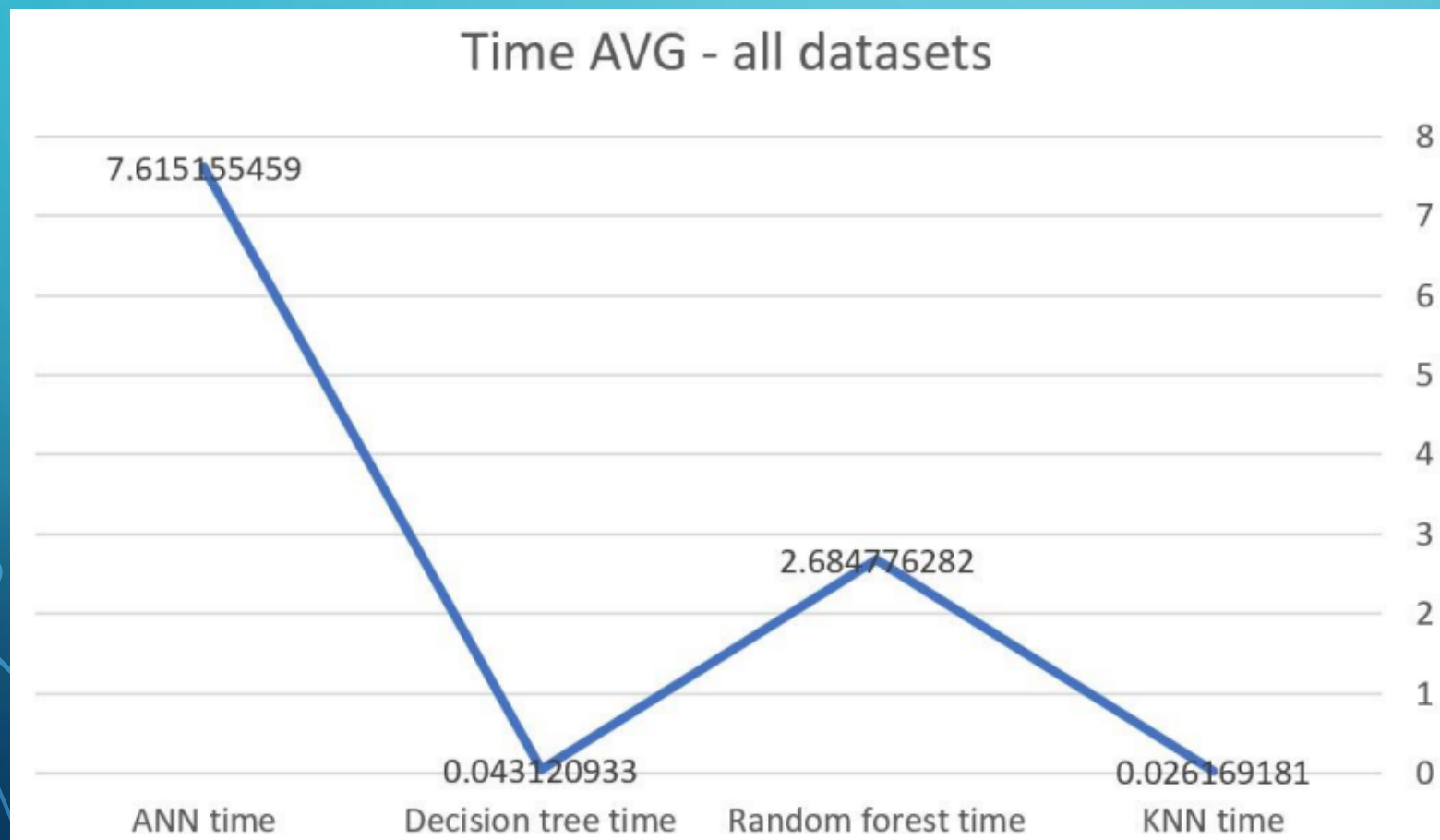
- עבור Artificial neural network בדקנו את הפרמטרים הבאים:

- סוג האקטיבציה ואת קצב הלמידה.

FLOW

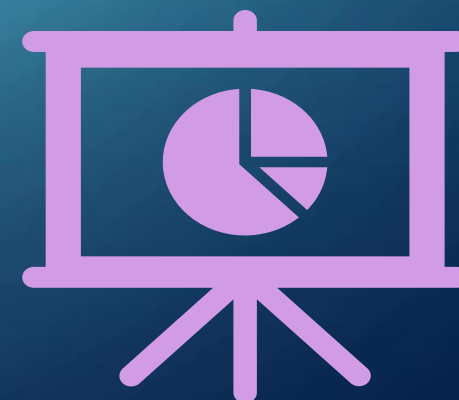


תוצאות

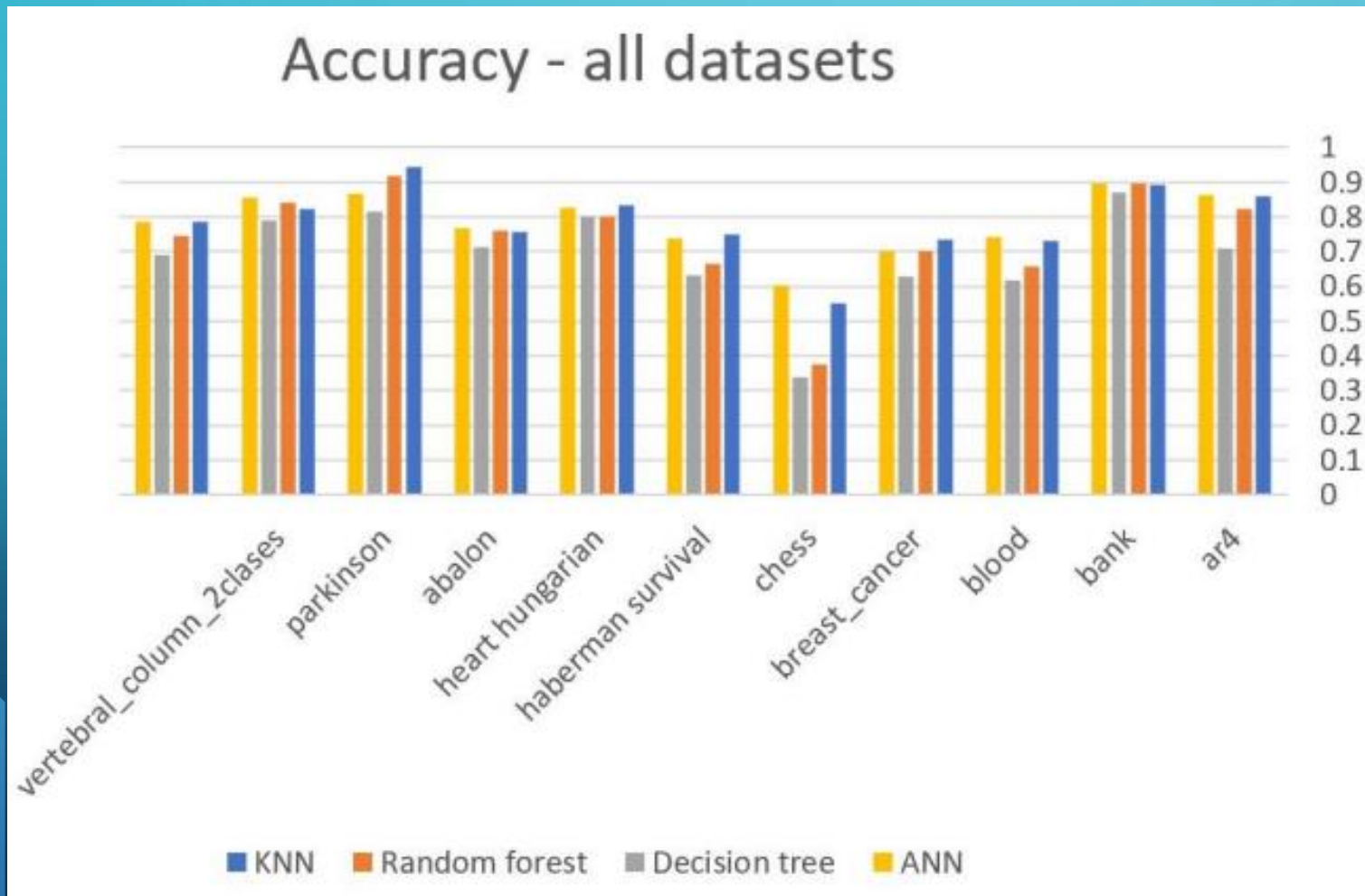


הזמן הממוצע של KNN הוא הנמוך ביותר.

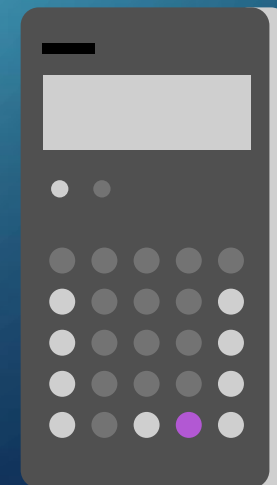
זה נובע מכך שהוא אלגוריתם פשוט, ושל ANN הוא הגבוה ביותר מכיוון שיש צורך בטעינה ושמירה של נתונים וזה גוזל זמן רב.



תוצאות



ה accuracy של chess הכי נמוך
כנראה מכיוון שהוא עם מספר
התכונות הגבוה ביותר.



הכרנו לראשונה את הרכיבים
השונים של תהליך הפיתוח

עיצבנו ותכננו את המערכת לפני
מסמכי אפיון ודרישות

פיתחנו את התוכנה שאפיינו
ועיצבנו עפ"י מודל השכבות

התנסינו לראשונה בכתיבת בדיקות
אוטומטיות מסוגים רבים

התנסינו לראשונה במחקר מודלים
שונים, חיזוי והסקת נתונים מהם

עבדנו לראשונה כצוות במקביליות
ובעבודת צוות על פרויקט גדול

לסיכום