TradeLog – מסמך אפיון וסיכום

פרויקט קורס · מסלול 3 (מוצר תוכנה ־ אפליקציית AI למשקיעים) · בן רובינוביץ׳, עידן, שחר · אוניברסיטת בן-גוריון · 2026 **האפליקציה החיה:** https://tradelog-peach.vercel.app · **מאגר הקוד:** `github.com/Benru1503/tradelog` · **מחברת המחקר:** `ml/tradelog_prediction.ipynb`

---

1. הבעיה הפיננסית

משקיעים פרטיים הסוחרים במניות, בקריפטו ובמט״ח ניצבים מול שלוש בעיות שמזינות זו את זו, ושתיים מהן TradeLog פותרת ישירות באמצעות AI:

1. **אין תמונת ביצועים מאוחדת וכנה.** הפוזיציות מפוזרות בין ברוקרים ובורסות; אמירה תמימה כמו ״החשבון עלה ב-10%״ מערבבת בין מיומנות מסחר לבין הפקדות ומשיכות, כך שהמשקיע אינו יכול לדעת אם הוא באמת משתפר.
2. **אין אינדיקציה צופה פני עתיד.** למשקיע עצמאי אין כלי נגיש שמעריך את ההסתברות שפוזיציה תעלה או תרד ביום או בשבוע הקרובים, עם רמת ביטחון כנה – ולא ״קנה/מכור״ של קופסה שחורה.
3. **אין משוב על *התנהגות* המסחר.** את התוצאות מכתיבים הרגלים – החזקת מפסידות זמן רב מדי, הגדלת פוזיציה אחרי הפסד, התעלמות מהיתרון האישי – אך שום כלי אינו חושף את הדפוסים האלה מתוך ההיסטוריה של הסוחר עצמו.

שתי יכולות ה-AI של TradeLog עונות ישירות על בעיות 2 ו-3: **Predict** (מודל מאומן לחיזוי כיוון, עם רמת ביטחון מוצהרת) ו-**Coach** (מודל שפה שהופך את היסטוריית המסחר של המשתמש לממצאים התנהגותיים). בעיה 1 נפתרת על ידי המוצר עצמו – היומן, הפוזיציות והתשואות המשוקללות בזמן ובכסף – ששתי יכולות ה-AI בנויות מעליו.

---

2. המוצר והארכיטקטורה הטכנולוגית

TradeLog היא אפליקציית Next.js 14 אחת (App Router, TypeScript) הנשענת על Supabase (+ Postgres, Google OAuth, אזור אירופה) ופרוסה ב-Vercel באותו אזור (פרנקפורט) שבו יושב מסד הנתונים – ללא שירות backend נפרד. הדפים הם React Server Components; כל פעולת כתיבה עוברת דרך server action או API route, כך שהדפדפן לעולם אינו ניגש למסד הנתונים ואינו מחזיק מפתחות API.

```
Browser (React)  ──HTTPS──►  Middleware (session refresh)  ──►  Server components / actions
                                                                  │
                                                  ┌───────────────┴────────────────┐
                                                  ▼                ▼               ▼
                                          Prisma → Postgres  Market-data     AI layer
                                                             router          (Predict, Coach)
```

יכולות הליבה: יומן עסקאות עם פוזיציות ועלות ממוצעת, ספר תזרים מזומנים שמזין תשואות TWR/MWR, אנליטיקה (עקומת הון, מפת חום ענפית ופילוח החזקות), רשימת מעקב, ו-Playground מבודד לסימולציות what-if ו-DCA. נתוני השוק (Finnhub לשערי מניות ומט״ח, Yahoo Finance להיסטוריית מחירים, CoinGecko לקריפטו) נשלפים בצד השרת בלבד, וכאשר ספק אינו זמין הממשק מציג `—` במקום לקרוס.

הנדסה ואבטחה. כל שאילתה מוגבלת לנתוני המשתמש שביצע אותה, RLS מופעל על כל טבלה, ולתפקיד הציבורי (anon) במסד הנתונים אין הרשאות כלל – נעילה שבוצעה במכוון אחרי שביקורת על הרשאות ברירת המחדל של Supabase מצאה טבלאות שהיו נגישות דרך מפתח ה-API הציבורי. מפתחות הספקים והמודלים קיימים אך ורק בתהליכי השרת – נבדק בסריקה של הקוד שנשלח בפועל לדפדפן. ההתחברות מוגבלת ל-Google בקוד האפליקציה ולא דרך הגדרה בלוח הבקרה, משום ש-API האימות נגיש ישירות עם המפתח הציבורי והגדרה עלולה להשתנות בלי שאיש ישים לב. נתיבי נתוני השוק וה-AI מוגבלים בקצב לכל משתמש. ה-CI חוסם כל דחיפה עד למעבר בדיקות lint, בדיקת טיפוסים, פורמט, בנייה

לייצור ו-**148 בדיקות יחידה** – ובהן בדיקות "וקטור זהב" שמצמידות את התאמת המודל בין Python ל-TypeScript (סעיף 3.1); בנוסף **11 בדיקות קצה-לקצה** ב-Playwright רצות מול מסד נתונים אמיתי, מחוץ ל-CI.

שני מודלי ה-AI

	Predict	Coach
משימה	כיוון המחיר ליום או לשבוע הקרוב, עם רמת ביטחון	סקירה התנהגותית של היסטוריית המסחר של המשתמש
טכנולוגיה	XGBoost (עצי החלטה עם gradient boosting), מאומן מראש	Gemini (מודל שפה), נקרא בזמן הבקשה
קלט	19 מאפיינים חסרי סקאלה מתוך סגירות ונפחים יומיים	סטטיסטיקות מחושבות מראש + קטעים מהערות היומן
פלט	הסתברות לעלייה, כיוון ורמת ביטחון	ממצאים מובְנים (JSON מאומת מול סכימה)
משרת	כל טיקר; היסטוריית המחירים נשלפת ללא מפתח API	את היסטוריית המסחר של המשתמש המחובר בלבד

---

3. מודלי ה-AI – פירוט טכני

3.1 Predict – חיזוי כיוון עם רמת ביטחון

מאומן על 14 נכסים סחירים (10 BTC, ETH, מניות גדולות, SPY, QQQ), סגירות יומיות מ-2020 ועד היום, בחלוקה כרונולוגית קפדנית: אימון עד 12/2024, ולידציה 01/2025–09/2025, וחלון **בדיקה** מ-10/2025 שהמודל לא ראה מעולם – לא באימון, לא בעצירה המוקדמת ולא בבחירת הסף. שני מסווגי **d1** – XGBoost (הנר היומי הבא) ו-**w1** (חמישה נרות קדימה) – מאומנים על 19 מאפיינים חסרי סקאלה (תשואות לוגריתמיות, יחסי ממוצעים נעים, RSI בשיטת Cutler, תנודתיות מתגלגלת, מרחק מטווח 20 הנרות האחרונים, ציון-Z לנפח, קידוד לוח שנה וסימון סוג הנכס). המודל שבייצור אינו כולל נתונים חלופיים (on-chain, מאקרו, סנטימנט) – עליו לדרג כל טיקר מתוך היסטוריית מחירים הנשלפת ללא מפתח API.

האימון מתבצע ב-Python וההסקה ב-TypeScript, בתוך שרת ה-Next.js, והשניים חייבים להסכים במדויק. לכן כל מאפיין מיושם בלולאות פשוטות ובחלון היסטוריה סופי (עד 50 נרות), ולא באמצעות אינדיקטורים מוכנים מספריות או החלקה רקורסיבית (EMA אמיתי, RSI לפי Wilder) שערכם תלוי בכמות ההיסטוריה שנשלפה במקרה. שלב האימון מייצא את המודל כ-JSON של העצים, יחד עם היסט (intercept) **נמדד** אמפירית (הייצוג הפנימי של `base_score` ב-XGBoost אינו נחשב אמין); מעריך ב-TypeScript עובר על אותם עצים עם השוואות סף ב- `float32` , בהתאמה מלאה למנוע החיזוי של XGBoost. בדיקות "וקטור זהב" מצמידות את המאפיינים ל-9 ספרות ואת ההסתברויות ל-8 ספרות מול המודל המאומן עצמו, כך ששינוי נוסחה בצד אחד בלבד מפיל את מערך הבדיקות מיד.

רמת הביטחון, בהגדרה מדויקת: המודל מחזיר `p(up)` , ההסתברות למחלקה "סוגר גבוה יותר בתום האופק". הכיוון הוא `UP` אם `p(up) ≥ 0.5` , ואחרת `DOWN` ; רמת הביטחון המוצגת היא `max(p(up), 1−p(up))` – כלומר עד כמה ההסתברות של המודל רחוקה מהטלת מטבע. זוהי ההסתברות הגולמית והבלתי מכוילת של המודל: אינדיקציית ביטחון כנה ואמיתית, אך כזו שלא אומתה מול מחקר כיול (ראו סעיף 5).

כל חיזוי נשמר עבור המשתמש, ובהמשך מוכרע מול שער עדכני לתוצאת HIT/MISS – כך שהרקורד ההיסטורי בפועל של האפליקציה גלוי בממשק, ולא רק מספרי backtest.

3.2 Coach – ניתוח התנהגותי באמצעות מודל שפה

Coach עונה על שאלה אחרת: לא "מה השוק יעשה", אלא "מה ההיסטוריה של הסוחר עצמו מספרת על אופן המסחר שלו". כל מספר שהמודל יכול לצטט – שיעור ההצלחה, אסימטריית זמני ההחזקה בין עסקאות מרוויחות למפסידות, יחס

הרווח הממוצע להפסד הממוצע (payoff ratio), שיעור עסקאות הנקמה, משמעת בגודל הפוזיציה, וביצועים לפי תגית ולפי יום בשבוע – מחושב דטרמיניסטית ב-TypeScript מתוך שורות ה-Prisma של המשתמש, *לפני* שהמודל רואה את הנתונים. מודל השפה (`gemini-flash-latest`, Gemini) מקבל את גיליון העובדות המוגמר ומונחה לפרש אותו בלבד ולא לחשב מספרים חדשים; הפלט מוגבל לסכימת JSON קבועה (`responseSchema` של Gemini) ומאומת שוב, עצמאית, ב-Zod בדרך חזרה – כך שתשובה פגומה או הזויה נכשלת בקול רם במקום להיות מוצגת לצד רווח והפסד אמיתיים. גיליון העובדות המדויק שמאחורי כל דוח נשמר ומוצג בממשק, כך שכל טענה בדוח ניתנת לביקורת מול המספרים שיצרו אותה.

העיצוב הזה מכוון בדיוק לכשל האופייני של תוכן פיננסי שנוצר במודלי שפה: מודל שמתבקש "לנתח את העסקאות האלה" ימציא בביטחון זמן החזקה ממוצע, ואילו מודל שקיבל `avgHoldHoursLosers: 12.4` יכול רק לחזור עליו. תפעולית: דוח דורש לפחות 5 עסקאות סגורות, נשמר במטמון לפי גיבוב (hash) של גיליון העובדות (היסטוריה שלא השתנתה מחזירה את הדוח השמור במקום לפנות שוב למודל), וכל משתמש מוגבל ל-10 הפקות בכל 24 שעות – משום שמפתח API אחד משרת את כל ההתקנה.

---

4. תוצאות

4.1 Predict – המודל שבייצור (חלון בדיקה 10/2025‑07/2026, 2,900‑2,956 שורות מוחזקות בצד)

אופק	AUC בבדיקה	דיוק בבדיקה	שיעור הבסיס ("עולה" תמיד)
d1 (יום הבא)	0.531	51.7%	51.8%
w1 (שבוע הבא)	0.527	51.7%	51.6%

הדיוק נמצא בגובה קו הבסיס הנאיבי ואף מעט מתחתיו; AUC מעט מעל 0.5 מעיד על אות **דירוג** קלוש, ולא על דיוק כיווני אמין – זהו המצב הכן של חיזוי כיוון יומי, וכך בדיוק הוא מוצג גם בממשק המוצר.

Backtest (לונג כאשר `p(up) ≥ 0.55`, אחרת מחוץ לשוק; עמלה של 10 נקודות בסיס בכל שינוי פוזיציה), המודל שבייצור:

נכס	אסטרטגיה	קנייה והחזקה	ימים בשוק	דרודאון מרבי	Sharpe
BTC	**+9.7%**	−46.1%	29%	−14.9%	0.54
AAPL	−5.8%	+30.5%	27%	−15.3%	−0.40
SPY	−0.1%	+12.3%	54%	−6.6%	0.04

תוצאת ה-BTC נראית חזקה, ועמודת הדרודאון מסבירה מדוע: המנגנון הוא הימנעות מירידות – המודל ישב מחוץ לשוק ברוב הצניחה שהורידה את "קנייה והחזקה" ב-46% – ולא כישרון כיווני. בשני הנכסים שעלו, היציאה מהשוק עלתה כסף אמיתי. מדובר באפקט תלוי-משטר, לא ביתרון כללי.

מחברת המחקר (`ml/tradelog_prediction.ipynb`) משווה באופן עצמאי שלושה מודלים על אותה חלוקה נטולת דליפה – XGB-lite (מאפייני הייצור בלבד), XGB-full (בתוספת נתוני on-chain ומאקרו) ו-LSTM:

מודל	AUC בבדיקה	דיוק	שיעור הבסיס
XGB-lite	0.512	51.8%	51.8%
XGB-full	0.494	51.6%	51.8%
LSTM	0.500	51.3%	51.1%

שלושתם יושבים בקצה הרעש הסטטיסטי, והנתונים החלופיים לא סייעו. **עד כמה ה-backtest שביר – נמדד בשתי דרכים.** המחברת מריצה את אותם כללי backtest על המודל החלופי שלה (XGB-full) ועל אותו חלון מחירים, ו-BTC יוצא 7.0%– במקום 9.7%+ – בעוד ש"קנייה והחזקה" יוצאת זהה בשתי ההרצות (46.1%–, עד נקודת הבסיס), מה שמוכיח שנתוני המחירים זהים ושכל הפער נובע מבחירת המודל. בדיקת רגישות העמלות מצביעה לאותו כיוון: אותה הרצה נותנת 1.7%– ב-0 נקודות בסיס, 7.0%– ב-10, ו-14.3%– ב-25. לכן התשובה הכנה לשאלת האלפא של הקורס היא זו של המחברת: **אין אלפא חיובית עקבית אחרי עמלות.** ה-9.7%+ הוא דגימה אחת מתוך התפלגות של גרסאות מודל והנחות עלות; האפקט היחיד שניתן להגן עליו הוא הימנעות מירידות בשוק יורד – וכך זה נאמר במפורש הן במחברת והן בתיעוד המוצר.

4.2 Coach – התוצאה היא ערובה ארכיטקטונית, לא מדד דיוק

Coach אינו מודל חיזוי, ולכן אין לו מדד דיוק לדווח עליו. ה"תוצאה" שלו היא הערובה שהארכיטקטורה אוכפת: כל מספר שמצוטט בדוח ניתן למעקב עד למספר אמיתי שחושב מראש, ומאומת מול הסכימה בכל תגובה. מה שכן ניתן לבדוק הוא השכבה הזו – והיא נבדקת: חישוב העובדות מכוסה במלואו בבדיקות יחידה (טיפול בעסקאות איזון, החרגת עסקאות שנמחקו, הבחנה בין null לאפס, תקרות גודל למטען), לצד בדיקה שמוודאת שסכימת התגובה של Gemini ומאמת ה-Zod מתארים בדיוק את אותו מבנה. בפועל, זה חושף באופן אמין דפוסים ממשיים וספציפיים – למשל יחס זמני החזקה שמראה שמפסידות הוחזקו פי כמה יותר זמן ממרוויחות, או אשכול של עסקאות נקמה עם שיעור הצלחה מתחת לקו הבסיס – במקום עצות גנריות.

---

5. מגבלות וסיכונים (Caveats)

Predict:

- **חלון בדיקה יחיד** (9 חודשים, משטר שוק אחד). הערכת walk-forward על פני מספר חלונות היא הצעד הנכון הבא לפני שנותנים ליתרון אמון נוסף.
- **הסף (0.55) נבחר מראש**, לא עבר אופטימיזציה, וההסתברויות **אינן מכוילות** – מספר הביטחון כן, אך לא אומת מול עקומת אמינות.
- **רגישות לבחירת המודל**: כפי שמוצג בסעיף 4.1, החלפת מערך המאפיינים שבייצור בגרסת הנתונים החלופיים של המחברת הופכת את סימן ה-backtest של BTC על אותם מחירים בדיוק. כל מספר backtest בודד צריך להיקרא עם סרגל השגיאה הזה בראש.
- **הסטת התפלגות בין אימון להסקה**: הקריפטו מאומן על `BTC-USD` מ-Yahoo אך מוגש מ-CoinGecko בזמן ההסקה; שערי הסגירה בין הזירות נבדלים בעשרות נקודות בסיס (המאפיינים חסרי הסקאלה מצמצמים את ההשפעה אך אינם מבטלים אותה).
- **ההכרעה מבוססת לוח שנה, בעוד התיוג מבוסס נרות**: חיזוי מוכרע 24 שעות (D1) או 7 ימים (W1) לאחר שנוצר – קירוב לאופק הנרות שעליו המודל אומן; תוצאה שטוחה בדיוק נספרת כ-MISS.
- **הטיית סל**: כל 14 הנכסים הם שמות גדולים ונזילים ששרדו עד 2026 – הטיית שרידות מתונה.
- תלוי בנקודת הקצה הלא רשמית (אך יציבה בפועל) של Yahoo Finance; בכשל, הממשק מציג הודעת "היסטוריה אינה זמינה" סטנדרטית במקום לקרוס.

Coach:

- פועל על **הרמה החינמית** של Google AI Studio – מכסות הבקשות הן לפי חשבון ומשותפות לכל משתמשי ההתקנה (ומכאן תקרת 10 הדוחות ביום), ותנאי הרמה החינמית של Google מתירים שימוש בתוכן שנשלח לשיפור מוצריהם (הרמה בתשלום אינה מתירה זאת). הערות היומן ומדדי המסחר נשלחים בכל דוח; עמוד הפרטיות מציין זאת, והדרך להימנע מכך היא פשוט לא להשתמש ביכולת.
- הקריאה הרשתית עצמה נבדקת ידנית ואינה מכוסה במערך הבדיקות האוטומטי (שכבת חישוב העובדות הדטרמיניסטית מכוסה במלואה).

- הערות יומן חופשיות זורמות אל ה-prompt; הנחיית המערכת מחייבת את המודל להתייחס אליהן כנתונים בלבד, והפלט מוגבל לסכימה וללא שימוש בכלים – אך הזרקת הנחיות (prompt injection) היא שיקול עיצובי, לא סיכון שבוטל מתמטית.

---

6. עבודה עתידית

- **אימון מחדש בשיטת walk-forward** על בסיס קבוע, במקום תמונת המצב הקפואה מ-07/2026.
- **מחקר כיול הסתברויות** (עקומות אמינות), כך שרמת הביטחון המוצגת תהיה מאומתת ולא רק מחושבת בכנות.
- **מאפייני מאקרו ללא מפתח** (SPX/VIX/DXY) **API** בזמן ההגשה – תוספת זולה שאינה שוברת את אילוץ "ללא מפתח".
- **התראות מודעות-פוזיציה** – הודעה למשתמש כשתחזית המודל מתהפכת עבור טיקר שהוא מחזיק.
- **יכולת AI נוספת בבחינה**: עוזר מבוסס-מסמכים (שאלות ותשובות עם אחזור מדוחות חברות) ככלי מחקר משלים ארוך-טווח, לצד Predict קצר-הטווח ו-Coach ההתנהגותי.