



# דוח פריסת הנתונים

השקעות בסיכון נמוך זה אצלנו-NexTrade

## תקציר

דוח זה מתאר את שלבי הפריסה וההטמעה של מערכת חיזוי אינטראקטיבית למדד S&P 500, שפותחה במסגרת פרויקט הגמר. הדוח כולל פירוט על סביבת הפיתוח, מבנה האתר, תהליך הטמעת המודל והתחזיות, אינטגרציה בין רכיבי המערכת, מנגנוני עדכון הנתונים ושיפורים עתידיים מוצעים. בנוסף, מובאות מסקנות מרכזיות מהתהליך והבנות שהתגבשו בשלב הסופי, לצד המלצות להמשך שיפור והרחבה של המערכת ככלי תומך החלטה למשקיעים.

מגישים – אריאל קריב ומשי בר

## החוג למערכות מידע

תוכן עניינים

2	סביבת הפיתוח
2	תיאור מבנה האתר
3	תיאור תהליך הפריסה
4	מנגנוני עדכון והזנת נתונים
4	אינטגרציה בין רכיבי המערכת
5	שיפורים עתידיים מוצעים

## סביבת הפיתוח

בפרויקט נעשה שימוש במגוון כלים, שפות ופלטפורמות, תוך התאמה לצרכים שונים בשלבי הפיתוח וההרצה:

- שפת Python שימשה לאיסוף הנתונים המקוריים ולאחר קניית המנוי ב-Alpha Vantage.
- Google Colab שימש לניקוי, בניית מאגר מאוחד, ניתוח רגשות ותחזית מדד, אימון מודלים מורכבים, עקב דרישות זיכרון גבוהות ויכולת ביצוע מקבילי.
- PyCharm שימש לשלבי הכנה ושליפת הנתונים בלבד.
- Visual Studio Code שימש לפיתוח ממשק המשתמש, עיצוב האתר ודינמיקה בין נתונים.
- HTML , CSS ו-JavaScript שימשו לבניית אתר האינטרנט ודשבורד ההדמיה.

## תיאור מבנה האתר

האתר פותח כסביבה אינטראקטיבית תומכת החלטה עבור משקיעים, וכולל שלושה עמודים מרכזיים:

- home.html עמוד פתיחה עם הסבר תמציתי על מטרות המערכת והצגת המבנה הכללי.



## החוג למערכות מידע

- analysis.html הצגה ויזואלית למשתמש לניתוחים סטטיסטיים על המדד בהתאם לתאריך, מחזק את ההבנה לחיזוי מהצג הטכני.
  - dashboard.html הדשבורד המרכזי בו מופיעים:
    - תחזית שבועית עדכנית מתוך- forecast\_latest.json
    - תחזיות היסטוריות קבצים- forecast\_week1.json, forecast\_week2.json, forecast\_week3.json
    - נתוני אמת של שוק המניות (actual\_data.json)
    - המלצות השקעה וניתוח רגשות (market\_dashboard\_data.json)
    - ניתוח רגשות מחדשות (news\_data.json)
    - הצגת גרפים של מגמות, תחזיות, סטטיסטיקות, והמלצות קונקרטיות.
- כל הנתונים נטענים בצורה דינמית ונשמרים בפורמט JSON לשם גמישות, נגישות וקלות עדכון.
- המערכת בנויה כך שניתן להרחיב אותה להזנת תחזיות חדשות בכל זמן, על ידי הפקת קבצי JSON נוספים והוספתם לתיקיית הפרויקט.

## תיאור תהליך הפריסה

- **רכישת גישה לנתונים:** בוצעה רכישת מפתח API בתשלום מ-Alpha Vantage על מנת לקבל נתונים היסטוריים מדויקים לאורך זמן.
- **ניקוי והכנה:** בוצע ניקוי של נתונים שגויים, סינון ימי מסחר חוקיים, המרת תאריכים וסטנדרטיזציה.

## החוג למערכות מידע

### • אימון המודל:

- נעשה שימוש במודל GRU
- המודל אומן על בסיס טווח נתונים של 30 יום והחזיר תחזית לשבעה ימים קדימה.
- המודל נשמר בקובץ `5final_snp500_model.h5` והסקיילר ב-`final_scaler.save`

### • הפקת התחזיות:

- התבצע ריצת Inference בשלושה שלבים: תחזית עתידית (`forecast_latest`) ותחזיות עבר להשוואה (`forecast_week1`, `forecast_week2`, `forecast_week3`)
- התחזיות נשמרו כקבצי JSON אשר משמשים את האתר.

### • יצירת דשבורד אינטראקטיבי:

- הדשבורד קורא קבצי JSON ומציג תחזיות באמצעות גרפים, טבלאות והמלצות.
  - הגרפים מציגים מגמות, השוואת תחזית מול מציאות, והסבר רגשי מבוסס אינדיקטורים.
- הסיבה לשימוש בקבצים מקומיים וקבועים בזמן, ולא בזמנים חיים, היא מגבלה תקציבית- שימוש בנתונים בזמן אמת דורש תשלום חודשי קבוע ל-API, ולכן יצרנו תחזיות בדיעבד עבור תאריכים מוגדרים מראש כדי לספק חוויית שימוש דינמית ככל האפשר למשתמש.

## מנגנוני עדכון והזנת נתונים

- הנתונים מוצגים מהקבצים:
  - forecast\_all.json חיבור של כל התחזיות לצורך השוואה.
  - market\_dashboard\_data.json חישוב RSI ו-MACD והמלצה BUY/SELL/HOLD
  - actual\_data.json נתונים אמיתיים להשוואת ביצוע.
- המערכת משתמשת בקוד JavaScript לטעינת הנתונים בצורה דינמית.
- הפלט מוצג לפי תאריכים חוקיים בלבד, תוך התחשבות בימי מסחר בפועל בלבד.
- חשוב לציין שהמשתמש לא מזין נתונים אשר צריכים להישמר לאורך כל השימוש שלו בכלי תומך ההחלטה שבנינו.

## אינטגרציה בין רכיבי המערכת

- הנתונים ההיסטוריים מוזנים לקוד הפיתוח אשר מבצע ניתוח תחזיתי.
- הקוד מייצר קבצי JSON עם התחזיות והמלצות.
- הדשבורד בנוי לקרוא את הקבצים דרך JavaScript ולשלב אותם בגרפים בזמן אמת.
- החיבור בין התחזית למציאות מתבצע באמצעות תאריכים תואמים ב־actual\_data.json.
- תחזיות מוצגות למשתמש עם הסבר טקסטואלי והמלצה בהתאם לשינוי במחיר הצפוי.

## שיפורים עתידיים מוצעים

### 1. אוטומציית חיזוי יומית

- בניית מנגנון אוטומטי שיריץ את המודל בכל בוקר וישמור קובץ תחזית חדש עם תאריך עדכני.

### 2. הוספת אפשרות השוואה בין מודלים

- פתיחת אפשרות למשתמש לבחור מודל ולהשוות בין תוצאות של GRU, LSTM או תחזית טכנית בלבד.

### 3. פריסת המערכת בענן (Heroku / Render)

- כיום המערכת רצה מקומית. ניתן להעלות את האתר לשרת חיצוני לצורך נגישות מכל מקום ושיתוף תוצאות בזמן אמת.

המערכת שהוקמה מדגימה אינטגרציה מתקדמת בין מידול נתונים, תובנות שוק והנגשה ויזואלית מקצועית, ומשמשת ככלי תומך החלטה מדויק ונוח עבור משקיעים.

## החוג למערכות מידע

לאורך הפרויקט ביצענו מסע מלא באתגרים, תובנות ולמידה מתמשכת- החל משלב הבנת הבעיה העסקית ועד לבניית מערכת אינטראקטיבית תומכת החלטה עבור משקיעים. אחת מאבני הדרך המרכזיות בפרויקט הייתה ההבנה שמטרתנו אינה רק לחזות את מדד S&P 500, אלא להנגיש את התחזיות והמשמעות שלן בצורה שתאפשר קבלת החלטות מושכלת - גם עבור מי שאינו מומחה טכני.

### אבני הדרך המרכזיות בפרויקט:

- **איסוף נתונים איכותי ובניית מאגר אחיד**- שילבנו נתוני מדד S&P 500 עם ניתוח רגשות מחדשות כלכליות שנמשכו ממנועי API. לצורך דיוק גבוה, בוצעה רכישת API בתשלום, ונעשו התאמות לנתוני מסחר בפועל בלבד (ימי פעילות).
- **בניית תשתית תחזית חזקה**- השוואה בין מודלים שונים (BERT, LSTM) הובילה לבחירה במודל GRU בזכות יציבות, ביצועים טובים ומהירות חישובית. בוצע תהליך הכנה מדוקדק של נתוני אימון, נרמול, חלוקה לרצפים, אימון ושמירת המודל והסקיילר לשימוש עתידי.
- **בניית דשבורד אינטראקטיבי**- השקענו בתכנון חוויית משתמש שתמחיש את התחזיות, תאפשר השוואה מול המציאות, תציג מגמות שוק, ותרכז נתונים חשובים בנקודת מבט אחת ברורה. הגרפים והתצוגות התבססו על קובצי JSON סטטיים - עקב מגבלות תקציב והיעדר גישה תמידית ל-API ח'.

### מסקנות ותובנות משמעותיות

- **אינטגרציה חכמה בין נתונים היא המפתח** ההחלטה לשלב בין נתונים כמותיים (טכניים) לבין רגשות מהחדשות הובילה ליצירת כלי עשיר יותר בתוכן, עם גמישות אנליטית גבוהה.





## **החוג למערכות מידע**

- **תחזית מול מציאות – לא תמיד צפויה** למדנו שישנם פערים בלתי נמנעים בין תחזית מדויקת ככל שתהיה לבין תנועות השוק בפועל, בעיקר בעקבות אירועים חיצוניים ( גיא-פוליטיים, רגולטוריים וכו').
- **הבנה עמוקה של המשתמש הסופי** במהלך הפרויקט הבנו שלצד המודלים המתקדמים, יש חשיבות עליונה לדרך שבה המידע מוצג. השקענו בהנגשה ברורה, הסברים טקסטואליים, והמלצות פשוטות, כדי לאפשר גם למשתמשים לא-טכנולוגיים להשתמש בכלי.

### **תובנות לשיפור לעתיד**

- **אוטומציה מלאה של ריצות תחזית** – כיום אנו תלויים בקבצים סטטיים. שילוב מנגנון חיזוי יומי יוכל לשפר משמעותית את הדינמיקה והשימושיות.
- **שדרוג מודל ניתוח רגשות** – קידוד רגשי מבוסס כללים שהשתמשנו בו מוגבל. שימוש במנועים מתקדמים (כמו FinBERT) יוכל להביא תובנות עשירות ומדויקות יותר.
- **העברת המערכת לענן** – הפיכת הדשבורד לזמין אונליין תאפשר גישה נוחה, שיתוף בין משתמשים, וניטור עדכני מכל מקום.

לסיכום, הצלחנו לבנות מערכת שלמה, המשקפת תהליך מדעי נתונים מלא – מאיסוף ועיבוד, דרך מידול ותחזית, ועד להצגת תובנות ויזואליות ברורות. הפרויקט לא רק העשיר אותנו בידע טכנולוגי, אלא גם חידד את החשיבות של תכנון נכון, התאמה לצרכי המשתמש, ושמירה על איזון בין מורכבות מדעית ונוחות שימוש.