# **תכנון ומימוש מערכת מסחר אלגוריתמית מתקדמת מרובת סוכנים לשוק החוזים העתידיים: ניתוח ארכיטקטוני ופונקציונלי**

## **I. תקציר מנהלים**

דוח זה מציג ניתוח מעמיק והצעה ארכיטקטונית למערכת מסחר אלגוריתמית מתוחכמת המיועדת לפעולה בשוק החוזים העתידיים. המערכת המוצעת משלבת באופן חדשני מספר רכיבים טכנולוגיים מתקדמים, ביניהם עיבוד נתוני נרות הייקן אשי, ניתוח צמתי נפח נמוך (LVN) מפרופיל שוק, ייצוג גיאומטרי של עסקאות באמצעות "צורות הרמוניות" (משולשים), זיהוי מצבי שוק באמצעות מנגנון Higher-Rank MMD, ולמידת חיזוק מרובת סוכנים (MARL) עם הטמעת קלט באמצעות רשתות נוירונים הנלמדת מקצה לקצה.

הדוח מפרט את הרכיבים האסטרטגיים, את תהליך עיבוד הנתונים והנדסת המאפיינים, את מבנה מטריצת הקלט המוצעת, את ארכיטקטורת רשת ההטמעה ואת מערך סוכני ה-MARL. דגש מיוחד מושם על האופן שבו הסוכנים לומדים לנהל סיכונים באופן דינמי, בדגש על התאמת גודל הפוזיציה בהתבסס על איכות האות הגיאומטרי ועוצמת תחנות ה-LVN.

המלצות הליבה של הדוח מתמקדות בבנייה מודולרית של המערכת, באימות קפדני של כל רכיב, ובחשיבות של תהליך אימון רובסטי. הפוטנציאל של מערכת זו לייצר אסטרטגיות מסחר אדפטיביות ועמידות בשווקים דינמיים הוא משמעותי, אך מימושו כרוך בהתמודדות עם אתגרים כגון הנדסת מאפיינים לאותות מבוססי-אירועים והבטחת למידה יציבה ואמינה של הסוכנים. הדוח מתווה נתיב לפיתוח מערכת זו, תוך התייחסות לשיקולים קריטיים הנוגעים למסחר חי.

## **II. צלילה לעומק אסטרטגיית המסחר האלגוריתמית**

בטרם ניגש לתכנון המערכת האוטומטית מבוססת סוכני למידת חיזוק (MARL), חיוני לנתח לעומק את הלוגיקה האסטרטגית העומדת בבסיסה. הבנה זו של רכיבי הליבה, האינדיקטורים, טווחי הזמן והאינטראקציות ביניהם, מהווה תנאי הכרחי לבניית מערכת אינטליגנטית שתדע לפרש את השוק ולפעול בו בצורה מיטבית. האסטרטגיה המוצעת נשענת על אימות רב-שכבתי, המשלב ניתוחי מומנטום, רגרסיה, חוסר איזון במחיר ומבנה שוק, וזאת על מנת לסנן רעשים ולהגביר את איכות אותות הכניסה.

### **A. מסגרת מרובת אינדיקטורים: MLMI, NW-RQK, FVG, MMD Market Regime, ו-LVN**

ליבת האסטרטגיה מורכבת ממספר אינדיקטורים וניתוחים הפועלים על נרות הייקן אשי בטווחי זמן שונים, ומספקים יחד תמונה רב-ממדית של השוק.

* **Machine Learning Momentum Index (MLMI) על נרות הייקן אשי של 30 דקות:**
  + **תיאור:** אינדיקטור ה-MLMI, כפי שמפורט במקורות 1, הוא מתנד מתקדם הפועל על גרף הייקן אשי של 30 דקות. הוא משלב עקרונות מומנטום מסורתיים (המבוססים על RSI וממוצעים נעים משוקללים - WMA) עם אלגוריתם k-Nearest Neighbors (k-NN) במטרה לייצר תחזיות מומנטום אדפטיביות. האות המרכזי לאסטרטגיה נוצר בעת חצייה של קו התחזית של MLMI עם ה-WMA שלו.1 פרמטרים מרכזיים המשפיעים על התנהגותו הם 'Prediction Data (k)' ו-'Trend length'.1 המשתמש הביע העדפה לשימוש בפרמטרי ברירת מחדל לשם שמירה על רובסטיות 1, נקודה המצוינת גם ב-.1
  + *פרשנות אדפטיבית של אותות מפרמטרים קבועים:* השימוש בפרמטרי ברירת מחדל עבור MLMI, מתוך שאיפה לרובסטיות, עשוי לוותר על שיפורי ביצועים פוטנציאליים הניתנים להשגה באמצעות כוונון ספציפי לנכס או למצב שוק.1 עם זאת, סוכן ה-MARL, על ידי התבוננות בפלט מנוע זיהוי מצבי השוק (MMD) ומאפיינים הקשריים אחרים, יכול ללמוד *לפרש* את האותות מ-MLMI בעל פרמטרים קבועים באופן שונה תחת תנאים משתנים. הדבר מאפשר הסתגלות משתמעת מבלי לשנות את האינדיקטור עצמו, מה שעולה בקנה אחד עם מטרת המשתמש לרובסטיות וההכרה בחוסר המעשיות של התאמת פרמטרים ידנית בזמן אמת.
    1. המשתמש מציין שימוש בפרמטרי ברירת מחדל של MLMI לשם רובסטיות.1
    2. מקור 1 מציין כי ביצועים מיטביים תלויים לרוב בכוונון פרמטרים, אך ניסיונו של המשתמש במסחר ידני מצביע על כך שפרמטרים קבועים מעשיים יותר עבור קו בסיס.
    3. תפקידו המרכזי של סוכן ה-MARL הוא ללמוד מדיניות אופטימלית (מיפוי ממצבים לפעולות).
    4. אם פרמטרי האינדיקטור קבועים, ייצוג המצב של הסוכן יכלול את מצב השוק כפי שזוהה על ידי MMD ונתונים הקשריים נוספים.
    5. לפיכך, הסוכן יכול ללמוד שאות חצייה ספציפי של MLMI עשוי להצביע על עסקה בעלת הסתברות גבוהה במצב שוק A, אך על עסקה בעלת הסתברות נמוכה במצב שוק B, ובכך להתאים את תגובתו לאות הקבוע בהתבסס על ההקשר השוקי הרחב יותר. הדבר משיג צורה של הסתגלות ברמה גבוהה יותר.
* **Nadaraya-Watson Rational Quadratic Kernel Regression (NW-RQK) על נרות הייקן אשי של 30 דקות:**
  + **תיאור:** אינדיקטור זה, הפועל גם הוא על גרף הייקן אשי של 30 דקות, משתמש ברגרסיית ליבה לא-פרמטרית עם ליבת Rational Quadratic ליצירת עקומת התאמה מיטבית למחיר.1 תכונה קריטית היא היותו "non-repainting" (אינו מצייר מחדש ערכים היסטוריים), החיונית לאמינות בדיקות לאחור וביצוע חי.1 אות המסחר נגזר כאשר עקומת הרגרסיה משנה את כיוונה, ומצביעה על מעבר לסנטימנט שורי או דובי.1 פרמטרים מרכזיים כוללים 'Bandwidth' ו-'Relative Weighting Parameter (Alpha)' 1, שגם עבורם המשתמש מעדיף את ערכי ברירת המחדל.
  + *חשיבות תכונת ה-"Non-Repainting" לבדיקות לאחור אמינות:* תכונת ה-"non-repainting" של NW-RQK 1 מהווה אבן יסוד לבניית אמון בביצועים ההיסטוריים של האסטרטגיה. אינדיקטורים "מציירים מחדש" (repainting), המשנים את ערכיהם ההיסטוריים עם קבלת נתונים חדשים, עלולים להוביל לתוצאות בדיקה לאחור אופטימיות מדי, שאינן משקפות את הביצועים שהיו מתקבלים במסחר חי.
    1. אינדיקטור NW-RQK מצוין במפורש כ-"non-repainting".1
    2. אינדיקטורים "מציירים מחדש" יוצרים סוג של הטיית מבט לעתיד (look-ahead bias) בבדיקות לאחור, שכן החלטות נראות כאילו התקבלו על סמך מידע שלא היה זמין באמת באותה נקודה היסטורית.
    3. המערכת של המשתמש מיועדת למסחר חי בכסף אמיתי (שאילתת משתמש1).
    4. לפיכך, אמינות הביצועים בבדיקות לאחור היא קריטית ביותר לאימות האסטרטגיה לפני פריסתה.
    5. תכונת ה-"non-repainting" של NW-RQK (והצורך להבטיח שגם מימושי MLMI ו-FVG יהיו כאלה) מחזקת משמעותית את תוקפם של מדדי ביצועים הנגזרים מסימולציות היסטוריות.
* **Fair Value Gap (FVG) - LuxAlgo על גרף של 5 דקות (הייקן אשי או סטנדרטי - טעון הבהרה):**
  + **תיאור:** אינדיקטור FVG, המסופק על ידי LuxAlgo, מזהה חוסר איזון בין היצע לביקוש על גרף של 5 דקות (שאילתת משתמש1 – יש לציין שהאסטרטגיה המקורית ב-1 השתמשה ב-FVG של דקה). האות המרכזי לאסטרטגיה הוא "מיטיגציה" (mitigation) של FVG, כאשר המחיר חוזר לאזור הפער וממלא אותו, לפחות חלקית.1 פרמטר 'Threshold %' מסייע בסינון גודל הפערים הנבחנים.1
  + *השפעת שינוי מסגרת הזמן של FVG על תזמון האות וגיאומטריית המשולש:* המעבר של המשתמש לניתוח FVG במסגרת זמן של 5 דקות (שאילתת משתמש), לעומת דקה כפי שתואר בניתוח האסטרטגיה הראשוני 1, צפוי להפחית את רעש האותות. עם זאת, שינוי זה גם יעכב מטבעו את אות המיטיגציה של FVG ביחס לאותות האינדיקטורים המהירים יותר הפועלים על 30 דקות (MLMI/NW-RQK), מה שישנה את המרווחים הזמניים ביניהם, וכתוצאה מכך, את המאפיינים הגיאומטריים של "המשולשים ההרמוניים".
    1. שאילתת המשתמש מציינת ניתוח FVG ושרטוט משולש הרמוני על גרף של 5 דקות.
    2. מסמך האסטרטגיה הראשוני 1 התייחס למיטיגציית FVG במסגרת זמן של דקה.
    3. אותות במסגרות זמן גבוהות יותר (5 דקות לעומת דקה) הם בדרך כלל פחות רועשים אך מופיעים בעיכוב רב יותר.
    4. אות המיטיגציה של FVG משמש כאחד משלושת הקודקודים היוצרים את המשולש ההרמוני.
    5. עיכוב באות ה-FVG (בשל גרף 5 הדקות) ביחס לאותות MLMI/NW-RQK (הנגזרים מגרפים של 30 דקות, אך נקודת ההקרנה המדויקת שלהם על גרף 5 הדקות מגדירה את הקודקוד שלהם) ישנה את אורכי צלעות המשולש ואת זוויותיו.
    6. מאחר שהאסטרטגיה מסתמכת על התכונות הגיאומטריות של משולשים אלו להערכת איכות הכניסה והסיכון, שינוי זה במסגרת הזמן של FVG משפיע ישירות על מרחב המאפיינים שממנו ילמד סוכן ה-RL. המערכת חייבת להיות מכוילת או ללמוד להתחשב בגיאומטריות משתנות אלו.
* **MMD Market Regime (על נרות הייקן אשי של 30 דקות):**
  + **תיאור:** רכיב קריטי הוא מזהה מצבי השוק, אשר יפעל על נתוני הייקן אשי של 30 דקות. המשתמש הביע העדפה לשיטות מתקדמות כגון Higher-Rank MMD 1, הממנפות חתימות נתיבים (path signatures) ללכידת דינמיקות שוק מורכבות. הדבר מספק הקשר חיוני לפירוש אותות המסחר.
  + *מצב שוק MMD כווסת מרכזי להתנהגות אדפטיבית:* פלט מנוע ה-MMD, כאשר ישולב במטריצת הקלט של הסוכן, ישמש כמשתנה מותנה רב עוצמה. הוא מאפשר למערכת ה-MARL ללמוד תגובות נפרדות לדפוסי אינדיקטורים דומים לכאורה, אם הם מתרחשים תחת תנאי שוק גלובליים שונים (למשל, תנודתיות גבוהה לעומת נמוכה, שוק מגמתי לעומת שוק מדשדש). זוהי דרך ישירה להתמודד עם אי-הנייחות (non-stationarity) המאפיינת שווקים פיננסיים.1
    1. שווקים פיננסיים הם אי-נייחים; תכונותיהם הסטטיסטיות משתנות לאורך זמן.1
    2. אסטרטגיות מסחר בעלות חוקים קבועים מתקשות לעיתים קרובות משום שאינן יכולות להסתגל לשינויים אלו.
    3. מנוע ה-MMD מתוכנן לזהות שינויים בהתפלגות הבסיסית של נתוני השוק, ובכך לזהות ביעילות מצבי שוק שונים.1
    4. על ידי הזנת הפלט של מנוע ה-MMD (תווית מצב או ציון) לסוכני ה-RL, הסוכנים יכולים ללמוד מדיניות המותנית במצב השוק הנוכחי.
    5. הדבר מאפשר למערכת להפגין התנהגות רובסטית ואדפטיבית יותר, לדוגמה, על ידי נקיטת פעולות אגרסיביות יותר במצבי שוק חיוביים וזהירות רבה יותר במצבים שליליים, או על ידי תעדוף דפוסי סינרגיה שונים בהתאם למצב השוק.
* **Low Volume Nodes (LVN) מפרופיל שוק (על נתוני 30 דקות):**
  + **תיאור:** צמתי נפח נמוך (LVNs), המזוהים מניתוח פרופיל שוק על נתוני 30 דקות, מיוצגים כתחנות מחיר משמעותיות. האסטרטגיה מניחה שכאשר המחיר מגיע ל-LVN, פוטנציאל העסקה המוצלחת גדל, שכן השוק "גילה כוונה" להגיע לאזור זה.1 היבט קריטי הוא הערכת "עוצמת התחנה" – המבוססת על בדיקות היסטוריות כתמיכה/התנגדות והנפח שנסחר ב-LVN – המשפיעה ישירות על החלטות גודל הפוזיציה.1
  + *"עוצמת תחנת" LVN כגורם דינמי נלמד לקנה מידה של סיכון/תגמול:* הרעיון של "עוצמת תחנה" מכניס רכיב ניהול סיכונים דינמי ישירות לתוך האסטרטגיה. כימות עוצמה זו באופן אובייקטיבי הוא משימת הנדסת מאפיינים מרכזית. סוכן ה-MARL יצטרך ללמוד את המתאם בין מדדי עוצמה מכומתים אלו (ומאפיינים הקשריים אחרים) לבין גודל הפוזיציה האופטימלי, ובכך ללמוד למעשה כיצד לקבוע את קנה המידה של הסיכון והתגמול הפוטנציאלי בהתבסס על החשיבות המבנית הנתפסת של רמת המחיר.
    1. המשתמש מגדיר LVNs כ"תחנות" ומציין ש"עוצמתן" אמורה להשפיע על גודל הפוזיציה.1
    2. "עוצמה" נקבעת על ידי אינטראקציות מחיר היסטוריות ב-LVN (מספר בדיקות כתמיכה/התנגדות, נפח שנסחר בבדיקות קודמות).
    3. הדבר דורש שיטה אלגוריתמית שיטתית לכימות אינטראקציות היסטוריות אלו ל"ציון עוצמה" מספרי או למערך מאפיינים המייצגים עוצמה.
    4. מאפייני עוצמת LVN אלו הופכים לחלק ממצב הקלט של סוכן ה-RL.
    5. רשת המדיניות של הסוכן לומדת אז מיפוי ממצב זה (הכולל עוצמת LVN, מאפייני משולש הרמוני, מצב שוק וכו') לגודל פוזיציה אופטימלי.
    6. הדבר הופך רעיון מסחר איכותני (עוצמת רמה) לקלט מכומת עבור פונקציית ניהול סיכונים דינמית ונלמדת.

### **B. ייצוג גאומטרי חדשני של עסקאות: משולשים הרמוניים**

אבן יסוד באסטרטגיה המוצעת על ידי המשתמש היא הפרשנות של צירוף שלושת אותות האינדיקטורים המרכזיים (MLMI, NW-RQK, ומיטיגציית FVG) לא רק כרצף אירועים זמני, אלא כיצירה של "צורה הרמונית" – ספציפית, משולש גאומטרי. משולש זה נבנה ומנותח על גרף של 5 דקות.1

* **מיצוי מאפיינים גאומטריים:** לאחר ששלושת קודקודי המשולש מוגדרים על ידי נקודות הזמן והמחיר של כל אחד משלושת האותות, ניתן לחלץ מהם מאפיינים גאומטריים כמותיים. מאפיינים אלו יכולים לכלול, בין היתר, את אורכי צלעות המשולש, הזוויות הפנימיות שלו, יחסים בין צלעות או בין גבהים לבסיסים, ושטח המשולש.1
* **מטרה אסטרטגית:** מאפיינים גאומטריים אלו נועדו לשמש כמדד ל"איכות" או "תוקף" של מערך הכניסה הפוטנציאלי. ההנחה הבסיסית היא שתצורות גאומטריות מסוימות של משולשים אלו מתואמות עם הסתברות גבוהה יותר להצלחת העסקה, והערכה זו אמורה להשפיע ישירות על החלטות ניהול הסיכונים, ובפרט על קביעת גודל הפוזיציה.1
* *מאפיינים גאומטריים כקלטים מהונדסים מתקדמים – חשיבות האימות:* טרנספורמציה זו של רצף אותות זמני לייצוג מרחבי/גאומטרי מהווה סוג חדשני ביותר של הנדסת מאפיינים. האתגר העיקרי יהיה לאמת באופן אמפירי את המובהקות הסטטיסטית של מאפיינים גאומטריים אלו. ללא בדיקות קפדניות, קיים סיכון של התאמת יתר (overfitting) לדפוסים אקראיים בנתונים היסטוריים (הטיית כריית נתונים - data mining bias). נרמול של מאפיינים גאומטריים אלו (לדוגמה, זוויות הן מנורמלות מטבען, אך אורכי צלעות ידרשו התאמה לתנודתיות או לרמת המחירים) יהיה חיוני לפירוש עקבי על ידי רשת נוירונים או סוכן RL בהמשך הדרך.
  1. המשתמש מציע שיטה חדשנית: המרת שלוש נקודות אות (זמן-מחיר) למשולש ומיצוי תכונותיו הגיאומטריות (שאילתת משתמש1).
  2. תכונות אלו (זוויות, אורכי צלעות, שטח) הופכות למאפייני קלט חדשים למערכת.
  3. ההשערה היא שלמאפיינים גאומטריים אלו יש כוח ניבוי לגבי תוצאת העסקה או איכותה.
  4. השערה זו חייבת לעבור אימות סטטיסטי על נתוני מחוץ-לדגימה (out-of-sample) כדי להבטיח שהדפוסים אמיתיים ולא רק תוצרים של הנתונים ההיסטוריים הספציפיים ששימשו לפיתוח.
  5. מאפיינים גאומטריים כמו אורכי צלעות (הנמדדים ביחידות מחיר) או שטח (הנמדד ביחידות מחיר כפול זמן) ישתנו באופן משמעותי עם רמת המחירים של הנכס והתנודתיות שלו.
  6. כדי להפוך מאפיינים אלו לברי-השוואה בין מצבי שוק ונכסים שונים, נרמול הוא חיוני (לדוגמה, אורכי צלעות כמכפלה של ATR, זוויות הן כבר בלתי-תלויות סקאלה).
  7. סוכן ה-RL ילמד ממאפיינים גאומטריים מנורמלים אלו; אם הם חסרי כוח ניבוי אמיתי, הם יוסיפו רעש ומורכבות ללא תועלת.
* *השפעת פער מסגרות הזמן על הגדרת קודקודי המשולש:* אותות ה-MLMI וה-NW-RQK מקורם בנתוני הייקן אשי של 30 דקות, בעוד שאות מיטיגציית ה-FVG והמשולש עצמו מוגדרים על גרף של 5 דקות. משמעות הדבר היא שהרזולוציה הזמנית עבור שניים מקודקודי המשולש גסה יותר. השיטה המדויקת להקרנת אירוע אות של 30 דקות על נר ספציפי של 5 דקות (למשל, הנר הראשון של 5 דקות בתוך בר ה-30 דקות שבו התרחש האות, הנר האחרון, או נר בעל מאפיינים ספציפיים כמו נפח גבוה ביותר) תשפיע ישירות על גיאומטריית המשולש, וכתוצאה מכך, על המאפיינים הנגזרים.
  1. אותות MLMI ו-NW-RQK נוצרים עם סגירת נר הייקן אשי של 30 דקות.
  2. אות מיטיגציית FVG נוצר על נר של 5 דקות.
  3. המשולש ההרמוני נבנה על גרף של 5 דקות תוך שימוש בשלושת האותות הללו כקודקודים.
  4. נר בודד של 30 דקות מכיל שישה נרות של 5 דקות.
  5. נדרש חוק ברור: איזה מבין ששת נרות ה-5 דקות הללו (המתאימים לבר האות של 30 הדקות) יגדיר את קואורדינטת הזמן/מחיר עבור קודקוד ה-MLMI/NW-RQK? לדוגמה, שימוש בפתיחה של נר ה-5 דקות הראשון בתוך בר האות של 30 הדקות, או בסגירה של נר ה-5 דקות האחרון.
  6. בחירה זו אינה טריוויאלית, שכן היא תשנה את קואורדינטות הזמן והמחיר של שני קודקודים, ובכך תשנה את כל התכונות הגיאומטריות של המשולש (אורכי צלעות, זוויות, שטח).
  7. עקביות בחוק זה חיונית ליציבות המאפיינים הגיאומטריים. ניתן אף לבחון את השפעתם של חוקי הקרנה שונים באמצעות ניתוח רגישות.

### **C. סינרגיות מונעות-שוק ולוגיקת כניסה**

* **עקרון ליבה:** כניסה לעסקה נשקלת רק כאשר כל שלושת האינדיקטורים המרכזיים (MLMI, NW-RQK, מיטיגציית FVG) מספקים אותות מאשרים באותו כיוון.1 אימות רב-גורמי זה נועד לסנן רעשי שוק ולשפר את דיוק הכניסה.
* **"ארבע הסינרגיות":** המשתמש מבהיר כי אלו אינן ארבע אסטרטגיות נפרדות שהסוכן בוחר ביניהן. במקום זאת, הן מייצגות את ארבעת הרצפים הזמניים האפשריים והתקפים שבהם שלושת אותות האינדיקטורים יכולים להופיע ולהתיישר ליצירת מערך כניסה שלם.1 דינמיקת השוק היא זו שמכתיבה איזה רצף מתממש בפועל. הרצפים הם:
  1. MLMI ← NW-RQK ← מיטיגציית FVG
  2. MLMI ← מיטיגציית FVG ← NW-RQK
  3. NW-RQK ← מיטיגציית FVG ← MLMI
  4. NW-RQK ← MLMI ← מיטיגציית FVG
* **תפקיד סוכן ה-RL:** תפקידו של הסוכן אינו "לבחור" סינרגיה. מרגע שסינרגיה (כלומר, רצף תקף של שלושה אותות מתואמים) הושלמה ונוצר משולש הרמוני, על מדיניות הסוכן להחליט *כיצד* לפעול. החלטה זו כוללת את עצם הכניסה לעסקה, את גודל הפוזיציה ואת ניהול הסיכונים שלאחר מכן, כל זאת בהתבסס על קלט המצב המקיף (הכולל את המאפיינים הגיאומטריים של המשולש שהתממש, הקשר LVN ומצב השוק).1
* *סינרגיות כדפוסים נצפים, לא כבחירות של הסוכן:* הבהרה זו היא קריטית. ה"סינרגיות" הן דפוסים מתהווים של אישורי אותות. סוכן ה-RL הוא מזהה דפוסים ומקבל החלטות מתוחכם הלומד להגיב באופן מיטבי לדפוסים נצפים אלו, ולא בורר מתוך תפריט קבוע של אסטרטגיות. סדר הרצף הספציפי שיצר את המשולש יכול להוות בעצמו מאפיין התורם לתכונות הגיאומטריות או קלט קטגוריאלי מפורש לסוכן.
  1. המשתמש תיאר בתחילה "ארבע סינרגיות", מה שיכול היה להתפרש כארבע אסטרטגיות נפרדות.
  2. הבהרתו המאוחרת של המשתמש 1 מציינת שהשוק מכתיב את הרצף, והסוכן מגיב לאחר היווצרות הסינרגיה.
  3. משמעות הדבר היא שלסוכן אין פעולה כמו "הפעל סינרגיה 2". במקום זאת, המצב שלו כולל מידע כגון "דפוס סינרגיה 2 הושלם זה עתה".
  4. מאפייני המשולש ההרמוני (שקודקודיו מוגדרים על ידי תזמוני האותות של הסינרגיה הספציפית) מקודדים באופן משתמע את הרצף. לדוגמה, הפרשי הזמן בין הקודקודים t1​,t2​,t3​ מגדירים את צורת המשולש וגם איזה אות הגיע ראשון, שני ושלישי.
  5. מדיניות הסוכן π(action∣state) לומדת למפות את המצב (הכולל את המאפיינים הגיאומטריים המשקפים את הסינרגיה שהתממשה, הקשר LVN, מצב השוק וכו') לפעולות מסחר מיטביות.
  6. הדבר מפשט את מרחב הפעולות של הסוכן (הוא אינו צריך "לבחור סינרגיה") אך מעשיר את מרחב המצבים שלו במאפייני הסינרגיה הנצפית.

### **טבלה 1: סקירת רכיבי הליבה האסטרטגיים**

| **שם הרכיב** | **מסגרת זמן (מקור נתונים)** | **סיכום לוגיקה/חישוב ליבה** | **אות/מאפיין לאסטרטגיה** | **פרמטרים מרכזיים/הערות תצורה** |
| --- | --- | --- | --- | --- |
| MLMI (Machine Learning Momentum Index) | 30 דקות הייקן אשי | שילוב RSI ו-k-NN לתחזית מומנטום אדפטיבית | חציית קו תחזית MLMI עם ה-WMA שלו | 'Prediction Data (k)', 'Trend length'. ברירת מחדל מועדפת. |
| NW-RQK (Nadaraya-Watson Rational Quadratic Kernel Regression) | 30 דקות הייקן אשי | רגרסיית ליבה לא-פרמטרית ליצירת עקומת התאמה למחיר (non-repainting) | שינוי כיוון עקומת הרגרסיה (שורי/דובי) | 'Bandwidth', 'Relative Weighting Parameter (Alpha)'. ברירת מחדל מועדפת. |
| FVG (Fair Value Gap) - LuxAlgo | 5 דקות (הייקן אשי/סטנדרטי - טעון הבהרה) | זיהוי פערי מחיר המייצגים חוסר איזון | "מיטיגציה" של FVG (חזרת המחיר למילוי הפער) | 'Threshold %'. |
| MMD Market Regime | 30 דקות הייקן אשי | זיהוי שינויים התפלגותיים בנתוני שוק באמצעות Higher-Rank MMD וחתימות נתיבים | תווית מצב שוק (קטגוריאלית) או ציון MMD (רציף) | פרמטרי MMD (חלונות, ליבה, סף וכו' - יפורט בפרק IV) |
| LVN (Low Volume Nodes) | פרופיל שוק מנתוני 30 דקות | זיהוי אזורי מחיר עם נפח מסחר נמוך היסטורית | "עוצמת תחנה" של LVN (מבוסס בדיקות קודמות, נפח) | הגדרות סף לזיהוי LVN, פרמטרים לכימות עוצמה |
| משולש הרמוני | 5 דקות (נוצר משילוב אותות MLMI, NW-RQK, FVG) | ייצוג גיאומטרי של צירוף שלושת האותות | מאפיינים גאומטריים (זוויות, יחסי צלעות, שטח וכו') | כללי הקרנת אותות 30 דקות ל-5 דקות |

טבלה זו מספקת נקודת ייחוס ברורה ומרוכזת לכל אבני הבניין של האסטרטגיה, מסגרות הזמן שלהן, ואופן תרומתן של אותות או מאפיינים לתהליך קבלת ההחלטות. היא מהווה בסיס להבנת תכנון צנרת הנתונים ומטריצת הקלט בהמשך.

## **III. תשתית נתונים והנדסת מטריצת קלט**

חלק זה מפרט את המסלול הקריטי מהמרת נתוני שוק גולמיים ועד ליצירת מטריצת קלט מובנית ועשירת-מאפיינים, המוכנה להזנה לשכבת ההטמעה של הרשת הנוירונית. תהליך זה כולל עיבוד מקדים קפדני, חישוב מאפיינים, והרכבה מחושבת של סוגי נתונים מגוונים.

### **A. רכישת נתונים, עיבוד מקדים (הייקן אשי) וניהול**

* **מקור נתונים וגרנולריות:** המערכת תשתמש בנתוני חוזים עתידיים, כאשר המשתמש ציין את Rithmic כמקור אפשרי.1 הנתונים יעובדו באינטרוולים של 30 דקות ו-5 דקות.1 נדרשים נתונים היסטוריים של שנתיים לצורך אימון ראשוני ובדיקות לאחור.
* **המרת נרות הייקן אשי:** שלב עיבוד מקדים מרכזי הוא המרת נרות OHLCV סטנדרטיים לנרות הייקן אשי, הן עבור מסגרות הזמן של 30 דקות והן עבור 5 דקות.1
  + **נוסחאות חישוב:** HA-Close = (O+H+L+C)/4; HA-Open = (HA-Open קודם + HA-Close קודם)/2; HA-High = Max(H, HA-Open, HA-Close); HA-Low = Min(L, HA-Open, HA-Close).1
  + **רציונל:** המשתמש מציין שהאסטרטגיה מותאמת באופן מיטבי לנרות הייקן אשי, ונרות אלו מסייעים בהפחתת רעשים, החלקת מגמות, ומספקים ייצוג פשוט יותר של המידע לכל נר.1
* **שלמות וניקוי נתונים:** יש ליישם נהלים סטנדרטיים לטיפול בנקודות נתונים חסרות (למשל, אינטרפולציה, מילוי קדימה, אם כי פחות קריטי עם אופי המיצוע של הייקן אשי), וזיהוי/טיפול בחריגות (שוב, הייקן אשי עשוי להחליק חלק מהחריגות). הבטחת חותמות זמן מדויקות וסנכרון בין זרמי הנתונים של 30 ו-5 דקות היא חיונית.
* **אחסון וניהול נתונים:** נדרשת מערכת חזקה לאחסון נתונים מעובדים היסטוריים (נרות הייקן אשי, אינדיקטורים מחושבים, מאפיינים) ולגישה יעילה אליהם לצורכי אימון ובדיקות לאחור. עבור מסחר חי, תידרש צנרת נתונים בזמן אמת.
* *השלכות המרת הייקן אשי אוניברסלית על FVG/LVN:* ההחלטה להמיר *את כל* נתוני המחיר להייקן אשי *לפני* חישוב אינדיקטורים כמו FVG ופרופיל שוק (עבור LVN) היא משמעותית. דפוסי FVG סטנדרטיים 1 מוגדרים על בסיס ערכי High ו-Low גולמיים של נרות ספציפיים. באופן דומה, פרופיל שוק נבנה בדרך כלל על בסיס רמות מחיר גולמיות והנפח שנסחר ברמות בדידות אלו. חישובם על נתוני הייקן אשי יניב דפוסים ורמות שונים.
  + הדוח חייב להבהיר במפורש האם חישובי FVG ו-LVN יותאמו לנתוני הייקן אשי (מה שידרוש הגדרות מותאמות אישית) או האם הם יחושבו על נתוני OHLCV גולמיים ואז האותות/רמות שלהם ייושרו עם אותות מבוססי הייקן אשי מ-MLMI/NW-RQK/MMD. לבחירה זו השלכות עמוקות על הנדסת המאפיינים ועל מטריצת הקלט.
    1. המשתמש מציין המרת כל הנתונים להייקן אשי.1
    2. MLMI, NW-RQK, ו-MMD יפעלו על נתוני הייקן אשי אלו.
    3. FVG ו-LVN (מפרופיל שוק) הם גם רכיבי מפתח באסטרטגיה.
    4. זיהוי FVG סטנדרטי מסתמך על יחסים ספציפיים בין ערכי High/Low גולמיים של שלושה נרות עוקבים.1 ערכי High/Low של נרות הייקן אשי מחושבים אחרת ומחליקים ערכים אלו.
    5. פרופיל שוק נבנה על ידי צבירת נפח ברמות מחיר גולמיות בדידות. רמות המחירים של הייקן אשי הן ממוצעות ומוסטות.
    6. לכן, יישום ישיר של לוגיקת FVG/פרופיל שוק סטנדרטית על נתוני הייקן אשי יביא לאזורי FVG ו-LVN שונים מאלו שהיו מתקבלים מנתונים גולמיים.
    7. אם הכוונה היא להשתמש ב-FVG/LVN סטנדרטיים, יש לחשבם על נתונים גולמיים. לאחר מכן, יש למפות או ליישר את התרחשותם/רמותיהם בזמן עם האותות מבוססי הייקן אשי לצורך בניית מטריצת הקלט והמשולשים ההרמוניים.
    8. אם FVG/LVN יחושבו *ישירות* מנתוני הייקן אשי, ייתכן שיהיה צורך להתאים אישית את הגדרותיהם ופרשנותם, שכן הם לא יהיו עוד "סטנדרטיים". בחירה זו משפיעה על אופי המאפיינים שיוזנו לרשת הנוירונים.

### **B. תכנון מטריצת הקלט המאוחדת (NxF): סכמה, תוכן ושילוב מאפיינים**

* **מטרה:** ליצור מטריצת קלט סדרתית במבנה של NxF עבור שכבת ההטמעה של הרשת הנוירונית, כאשר N הוא אורך הרצף (מספר צעדי זמן היסטוריים) ו-F הוא מספר המאפיינים בכל צעד זמן. מטריצה זו חייבת לייצג באופן מקיף את מצב השוק בהתאם לאסטרטגיה.
* **גרנולריות צעד הזמן:** כל שורה במטריצת NxF תייצג ככל הנראה צעד זמן בדיד. מאחר שהמשולשים ההרמוניים משורטטים על גרף של 5 דקות, הגיוני שצעד הזמן הבסיסי של המטריצה יהיה 5 דקות.
* **אורך הרצף (N):** פרמטר זה מגדיר את חלון ההסתכלות לאחור. בחירת N היא קריטית ומשפיעה על יכולת המודל ללכוד תלויות זמניות. זהו היפר-פרמטר שיש לכייל (6 דנים באורך רצף עבור LSTMs/Transformers; 7 מזכיר 240 צעדי זמן לנתונים יומיים; 6 משתמש ב-15 עבור הכנסות רבעוניות). עבור נתוני 5 דקות, N יכול לנוע בין מספר שעות (למשל, N=24 עבור שעתיים) ועד לסשן מסחר מלא או יותר, בהתאם למגבלות חישוביות ולזיכרון הרצוי של המערכת.
* **הרכב וקטור המאפיינים (F מאפיינים לכל צעד זמן של 5 דקות):**
  1. **נתוני נרות הייקן אשי (5 דקות):** HA-Open, HA-High, HA-Low, HA-Close, Volume עבור אינטרוול 5 הדקות הנוכחי.
  2. **נתוני נרות הייקן אשי (הקשר של 30 דקות):** HA-Open, HA-High, HA-Low, HA-Close, Volume עבור בר הייקן אשי של 30 דקות שאליו שייך בר 5 הדקות הנוכחי. הדבר מספק הקשר רב-מסגרות זמן באופן ישיר.
  3. **ערכי אינדיקטורים (מיושרים לצעדי 5 דקות):**
     + MLMI (30 דקות HA): ערך נוכחי של קו MLMI, WMA של MLMI. (ערך מהבר האחרון הסגור של 30 דקות, נישא קדימה עבור ברי 5 הדקות העוקבים עד לעדכון הבא).
     + NW-RQK (30 דקות HA): ערך נוכחי של עקומת הרגרסיה, שיפוע/כיוון. (ערך מהבר האחרון הסגור של 30 דקות, נישא קדימה).
     + FVG (5 דקות, פוטנציאלית מבוסס HA): דגלים בינאריים לקיום FVG שורי/דובי, מרחק לגבולות FVG, סטטוס מיטיגציה.
  4. **מאפייני LVN (מיושרים לצעדי 5 דקות):**
     + קרבה (למשל, מרחק בטיקים או מכפלות ATR) ל-LVN המשמעותי הקרוב ביותר שזוהה בפרופיל של 30 דקות.
     + ציון "עוצמת תחנה" מכומת עבור LVN(s) רלוונטי(ים) אלו. (ערכים מעדכון פרופיל 30 הדקות האחרון, נישאים קדימה).
  5. **מאפייני משולש הרמוני (מבוססי-אירוע, מיושרים לצעדי 5 דקות):**
     + תכונות גיאומטריות (זוויות, אורכי צלעות מנורמלים, שטח) של המשולש ההרמוני *האחרון שהושלם*.
     + זמן שחלף (בברי 5 דקות) מאז היווצרות המשולש האחרון.
     + דגל בינארי: 1 אם משולש חדש הושלם בבר 5 הדקות הנוכחי, 0 אחרת.
  6. **פלט מצב שוק MMD (מיושר לצעדי 5 דקות):**
     + תווית מצב נוכחי (למשל, מקודדת one-hot) או ציון MMD רציף. (ערך מעדכון MMD של 30 הדקות האחרון, נישא קדימה).
* **נרמול וסקיילינג:** חיוני ליציבות וביצועי הרשת הנוירונית. כל אחד מ-F המאפיינים צריך לעבור סקיילינג מתאים (למשל, סטנדרטיזציית Z-score עבור תשואות או אינדיקטורים בלתי חסומים, סקיילינג Min-Max עבור אינדיקטורים חסומים כמו RSI או זוויות [0-180/0-PI]) על פני מערך נתוני האימון.8
* *מידע זמני היררכי בווקטור המאפיינים:* וקטור המאפיינים משלב מידע מרזולוציות זמן שונות (5 דקות בזמן אמת, 30 דקות הקשרי/מבוסס-אירוע). שכבת ההטמעה של הרשת הנוירונית תצטרך ללמוד לעבד מבנה זמני היררכי זה. מאפיינים מחישובי 30 דקות (MLMI, NW-RQK, LVN, MMD) יהיו קבועים למשך שש שורות עוקבות של 5 דקות במטריצה, ואז יתעדכנו. אופי קבוע-למקוטעין זה הוא בעצמו דפוס שהרשת הנוירונית יכולה ללמוד.
  1. למטריצת הקלט רזולוציית צעד זמן של 5 דקות.
  2. מספר תשומות מפתח (MLMI, NW-RQK, מאפייני LVN, מצב MMD) נגזרות מחישובים של 30 דקות.
  3. בעת אכלוס וקטור המאפיינים ברזולוציה של 5 דקות, הערכים עבור מאפיינים נגזרי-30-דקות אלו יישארו קבועים למשך שישה צעדי 5 דקות ואז יתעדכנו כאשר בר ה-30 דקות הבא ייסגר וחישובים חדשים יהיו זמינים.
  4. הדבר יוצר דפוס "בלוקי" או קבוע-למקוטעין עבור מאפיינים ספציפיים אלו בתוך מטריצת ה-NxF.
  5. שכבת ההטמעה של הרשת הנוירונית (למשל, LSTM או Transformer) תראה דפוסים אלו. עליה ללמוד שלמאפיינים אלו תדירות עדכון שונה והם מייצגים הקשר הנע לאט יותר בהשוואה לנתוני FVG או הייקן אשי של 5 דקות.
  6. תזמון היררכי זה הוא היבט חיוני של ניתוח רב-מסגרות-הזמן שהמשתמש שואף אליו.

### **C. טיפול במאפיינים מבוססי-אירועים וספורדיים בתוך רצפים באורך קבוע**

* **האתגר:** משולשים הרמוניים הם מונעי-אירועים; הם אינם נוצרים בכל אינטרוול של 5 דקות. גם מיטיגציות FVG הן אירועים. מטריצת ה-NxF, לעומת זאת, דורשת מערך עקבי של F מאפיינים בכל צעד זמן N.
* **אסטרטגיית ייצוג למשולשים הרמוניים:**
  + **דגל אירוע נוכחי:** מאפיין בינארי המציין האם משולש הרמוני *חדש* הושלם בצעד הזמן הנוכחי של 5 דקות.
  + **מאפייני האירוע האחרון:** אחסון התכונות הגיאומטריות (זוויות, אורכי צלעות מנורמלים וכו') של המשולש *האחרון שהושלם*. ערכים אלו נישאים קדימה עד להיווצרות משולש חדש.
  + **זמן מאז האירוע האחרון:** מאפיין מספרי הסופר את מספר ברי ה-5 דקות מאז הושלם המשולש ההרמוני האחרון. הדבר מסייע למודל להעריך את "טריות" או רלוונטיות המאפיינים הגיאומטריים הנישאים קדימה. מקורות 11 דנים בטיפול באירועים נדירים ובבחירת מודלים בהתבסס על תצפיות סדרות זמן, ומדגישים את הצורך בייצוג חזק.
* **אסטרטגיית ייצוג לאותות FVG:**
  + **סטטוס אזור FVG:** מאפיינים המציינים האם קיים אזור FVG שורי/דובי, המרחק לגבולות הפער, סטטוס מיטיגציה (למשל, לא התרחשה, התרחשה חלקית, התרחשה במלואה). מאפיינים אלו יכולים להתעדכן בכל צעד זמן של 5 דקות בהתבסס על תנועת המחיר ביחס לפערים מזוהים.
* *בחירת ייצוג לאירועים ספורדיים יכולה להשפיע משמעותית על הלמידה.* נשיאת נתוני אירוע ישנים קדימה עשויה לספק הקשר אך גם להכניס מידע לא רלוונטי אם לא מנוהלת עם מאפיין טריות. ערכי מילוי-מקום דורשים מהרשת הנוירונית ללמוד להתעלם מהם או לפרשם נכונה.
  1. משולשים הרמוניים הם מפתח אך אינם נוצרים בכל אינטרוול של 5 דקות.
  2. מטריצת הקלט NxF דורשת מערך עקבי של F מאפיינים בכל אחד מ-N צעדי הזמן.
  3. כיצד לייצג את מאפייני המשולש "הנוכחי" כאשר לא נוצר אחד חדש?
  4. אם משתמשים במאפייני המשולש *הידוע האחרון*, הסוכן מקבל קלט גיאומטרי רציף, אך הוא עשוי להיות מאירוע ישן ולא רלוונטי.
  5. הוספת "זמן מאז האירוע" מסייעת לסוכן להפחית את משקלם של נתונים גיאומטריים ישנים.
  6. לחלופין, ריפוד באפסים או שימוש בדגלים מיוחדים עבור "אין אירוע נוכחי" הופכים את המאפיינים הגיאומטריים לספורדיים באמת בווקטור הקלט, והרשת הנוירונית חייבת ללמוד להתמודד עם כך.
  7. החלטה זו משפיעה על מורכבות מה ששכבת ההטמעה וסוכן ה-RL צריכים ללמוד.

### **טבלה 2: סכמת מטריצת קלט מאוחדת (מבנה NxF רעיוני עבור צעד זמן בודד)**

| **קטגוריית מאפיין** | **שם מאפיין ספציפי** | **סוג נתונים** | **שיטת נרמול/סקיילינג מוצעת** |
| --- | --- | --- | --- |
| הייקן אשי 5 דק' | HA\_Open\_5m, HA\_High\_5m, HA\_Low\_5m, HA\_Close\_5m | רציף | אחוז שינוי / לוג-תשואה, ואז סטנדרטיזציה |
|  | HA\_Volume\_5m | רציף | סטנדרטיזציה או Min-Max (אם חסום) |
| הייקן אשי 30 דק' (הקשר) | HA\_Open\_30m, HA\_High\_30m, HA\_Low\_30m, HA\_Close\_30m | רציף | כנ"ל (ערך נישא קדימה) |
|  | HA\_Volume\_30m | רציף | כנ"ל (ערך נישא קדימה) |
| MLMI (מ-30 דק' HA) | MLMI\_Value, MLMI\_WMA\_Value | רציף | Min-Max (אם האינדיקטור חסום) או סטנדרטיזציה (ערך נישא קדימה) |
| NW-RQK (מ-30 דק' HA) | NWRQK\_Value, NWRQK\_Slope | רציף | סטנדרטיזציה (ערך נישא קדימה) |
| סטטוס FVG (מ-5 דק') | FVG\_Bull\_Active, FVG\_Bear\_Active | בינארי | - |
|  | FVG\_Distance\_Upper, FVG\_Distance\_Lower | רציף | נרמול לפי ATR או טווח מחירים |
|  | FVG\_Mitigation\_Status | קטגוריאלי (למשל, 0=אין, 1=חלקי, 2=מלא) | One-hot encoding |
| הקשר LVN (מפרופיל 30 דק') | LVN\_Nearest\_Distance, LVN\_Nearest\_Strength\_Score | רציף | נרמול לפי ATR; Min-Max לציון (ערך נישא קדימה) |
| גיאומטריית משולש הרמוני | Triangle\_Angle\_A, Triangle\_Angle\_B, Triangle\_Angle\_C | רציף (0-180) | Min-Max לטווח או [1-,1] |
|  | Triangle\_Side\_Ratio\_12, Triangle\_Side\_Ratio\_23 | רציף | סטנדרטיזציה (אם הטווח אינו מוגדר היטב) |
|  | Triangle\_Area\_Normalized | רציף | Min-Max לאחר נרמול ראשוני (למשל, חלקי ATR בריבוע) |
|  | Triangle\_Time\_Since\_Formation | שלם (מספר ברים) | Min-Max על פני חלון היסטורי סביר או לוג-טרנספורמציה |
|  | Triangle\_Newly\_Formed\_Flag | בינארי | - |
| מצב שוק MMD (מ-30 דק' HA) | MMD\_Regime\_Label (אם קטגוריאלי) | קטגוריאלי | One-hot encoding (ערך נישא קדימה) |
|  | MMD\_Score (אם רציף) | רציף | סטנדרטיזציה (ערך נישא קדימה) |

טבלה זו עונה ישירות על שאלת הליבה של המשתמש לגבי סכמת המטריצה. היא מספקת תוכנית קונקרטית להנדסת מאפיינים ובניית צנרת נתונים, ומבטיחה שכל הרכיבים האסטרטגיים שצוינו יילקחו בחשבון בקלט לרשת הנוירונים. היא מבהירה את הממדיות (F) של וקטור המאפיינים בכל צעד זמן.

## **IV. זיהוי מצבי שוק מתקדם באמצעות Higher-Rank MMD**

חלק זה מתמקד בתיאוריה ובמימוש המעשי של מנגנון זיהוי מצבי השוק הנבחר, המהווה רכיב קריטי במערכת המסחר האדפטיבית.

### **A. עקרונות ומימוש (חתימות נתיבים, שיטות ליבה)**

* **רציונל לבחירת MMD:** המערכת תשתמש במבחן Maximum Mean Discrepancy (MMD) הלא-פרמטרי לצורך השוואת התפלגויות של נתוני שוק.1 המשתמש הביע העדפה לגרסה מתקדמת, Higher-Rank MMD, בשל יכולתה ללכוד תלות בנתיב (path-dependency) ודינמיקות לא-מרקוביות באמצעות שימוש בחתימות נתיבים מחוספסים (rough path signatures).1
* **חתימות נתיבים (Path Signatures):** אלו הן ייצוגים מתמטיים של סדרות זמן הלוכדים את הגיאומטריה שלהן ואת סדר התנועות באופן שאינו תלוי בפרמטריזציה מחדש של הזמן.1 החתימות יופקו ממאפייני נרות הייקן אשי (למשל, לוג-תשואות, תנודתיות) במסגרת זמן של 30 דקות.1
* **שיטות ליבה (Kernel Methods):** ליבת חתימה (signature kernel) תשמש בשילוב עם MMD להשוואת התפלגויות של נתיבים במרחב הילברט עם ליבה משחזרת (RKHS).1
* **Higher-Rank MMD:** הרחבה זו של MMD משלבת מידע פילטרציה (התפתחות המידע לאורך זמן), מה שהופך אותה לרגישה יותר לשינויים עדינים בדינמיקת השוק.1
* **מימוש:** המימוש יתבסס על ספריות פייתון כגון issaz/signature-regime-detection ו-higherOrderKME.1 דיון בפרמטרים: גודלי חלונות, בחירת ליבה, רמת קיטוע חתימה, סף סטטיסטי.1
* **זיהוי מקוון (Online Detection):** השוואת חלון נתונים עדכני לחלון ייחוס היסטורי לצורך זיהוי שינויים בזמן אמת.1
* הבחירה ב-"Higher-Rank MMD" היא מתוחכמת ועתירת חישוב. יתרונה המרכזי הוא הרגישות למאפיינים תלויי-נתיב ולא-מרקוביים, המאפיינים שווקים פיננסיים. האתגר המעשי יהיה כיול הפרמטרים שלה לאיזון אופטימלי בין תגובתיות לתוצאות חיוביות שגויות (false positives), במיוחד בהינתן תשומות הייקן אשי המוחלקות.
  1. המשתמש מעוניין בזיהוי מצבי שוק מתקדם.
  2. Higher-Rank MMD עם חתימות נבחר בשל יתרונותיו התיאורטיים עם סדרות זמן פיננסיות.
  3. שיטה זו כוללת מתמטיקה מורכבת (נתיבים מחוספסים, RKHS).
  4. המימוש מסתמך על ספריות ייעודיות.25
  5. פרמטרים מרכזיים (גודלי חלונות, סוג ליבה, עומק חתימה, סף MMD) דורשים כיול קפדני.1
  6. החלקת הייקן אשי עשויה להקשות על זיהוי שינויים התפלגותיים עדינים, מה שעשוי לדרוש פרמטרי MMD רגישים יותר.
  7. האיזון הוא בין לכידת שינויי מצב אמיתיים במהירות לבין הימנעות מהפעלה על ידי רעש או תנודות קלות המוגברות על ידי הגדרות רגישות.
* פלט ה-MMD (ציון מרחק או תווית מצב) הוא סיכום בעל ממדיות נמוכה יחסית של השוואה התפלגותית מורכבת. סוכן ה-RL לא יראה את חישוב ה-MMD הגולמי אלא את תוצאתו, מה שמפשט את מרחב המצבים של הסוכן אך גם מפשיט את הניואנסים של השינוי ההתפלגותי.
  1. MMD משווה התפלגויות של חתימות נתיבים.
  2. התוצאה היא סטטיסטי MMD יחיד או תווית מצב נגזרת.
  3. ערך/תווית יחיד זה מוזן למטריצת הקלט הגדולה הרבה יותר עבור סוכן ה-RL.
  4. הסוכן לומד לקשר פלט MMD זה לפעולות אופטימליות.
  5. משמעות הדבר היא שהסוכן אינו לומד *מדוע* מצב השוק השתנה, אלא רק *ש*הוא השתנה וכיצד להגיב.
  6. זוהי הפשטה המקלה על משימת הסוכן אך מסתמכת במידה רבה על יכולתו של מנוע ה-MMD לזהות נכונה שינויים משמעותיים.

### **B. שילוב כמאפיין במטריצת הקלט**

* **פלט מנוע ה-MMD:** יכול להיות ציון MMD רציף (המרחק בין ההתפלגויות) או תווית מצב קטגוריאלית.1
* **הכללה במטריצת NxF:** פלט זה הופך לאחד מ-F המאפיינים בכל צעד זמן, ומספק הקשר שוק גלובלי לסוכן ה-RL.
* **תדירות עדכון:** חישוב ה-MMD יתבצע באופן תקופתי (למשל, כל 30 דקות, בהתאם לאינטרוול נרות הייקן אשי המשמש כקלט שלו), והפלט האחרון שלו ישמש לאכלוס מאפיין המצב במטריצת הקלט עבור אינטרוולי 5 הדקות העוקבים עד לעדכון ה-MMD הבא.
* אם פלט ה-MMD הוא ציון רציף, הוא עשוי לספק אינדיקציה מורכבת יותר לגבי עוצמת המעבר בין מצבים או רמת הביטחון בזיהוי, ששכבת ההטמעה של הרשת הנוירונית תוכל פוטנציאלית ללמוד לפרש. תווית קטגוריאלית פשוטה יותר אך מאבדת ניואנס זה.
  1. פלט MMD יכול להיות רציף (ציון) או קטגוריאלי (תווית).
  2. פלט זה הוא מאפיין עבור הרשת הנוירונית/סוכן ה-RL.
  3. ציון רציף עשוי לאפשר לסוכן ללמוד דקויות, למשל, "מצב השוק *מתחיל* להשתנות" לעומת "מצב השוק השתנה *באופן מובהק*".
  4. תווית קטגוריאלית מפשטת את הקלט אך מאלצת סיווג קשיח של מצב השוק.
  5. הבחירה משפיעה על מה ששכבת ההטמעה וסוכן ה-RL צריכים ללמוד ממאפיין ספציפי זה.

### **טבלה 3: תצורת מנוע זיהוי מצבי שוק (MMD)**

| **פרמטר** | **ערך נבחר / אסטרטגיה** | **רציונל / השפעה על הזיהוי** |
| --- | --- | --- |
| מאפייני קלט לחתימת נתיב | לוג-תשואות ותנודתיות מחושבת (מנרות הייקן אשי של 30 דקות) | לכידת דינמיקת מחיר ותנודתיות כבסיס לזיהוי שינויים התפלגותיים. |
| רמת קיטוע חתימה (M) | למשל, 2 או 3 | קובעת את עומק לכידת התלות הסדרתית; ערכים גבוהים יותר לוכדים יותר מידע אך מגדילים מורכבות חישובית. |
| סוג ליבה (Kernel) עבור MMD | למשל, ליבה גאוסיאנית (RBF) על מרחב החתימות | בחירה נפוצה וגמישה; דורשת כיול פרמטרי ליבה (למשל, רוחב פס σ). |
| גודל חלון ייחוס היסטורי | למשל, 60-120 נרות של 30 דקות (30-60 ימי מסחר) | צריך להיות ארוך מספיק לייצג מצב יציב, אך לא ארוך מדי כדי לא לכלול מצבים ישנים מדי. |
| גודל חלון נוכחי (לבדיקה) | למשל, 20-30 נרות של 30 דקות (10-15 ימי מסחר) | צריך להיות רגיש מספיק לשינויים, אך לא קצר מדי כדי למנוע רעש. |
| סף MMD סטטיסטי לזיהוי שינוי | ייקבע אמפירית או באמצעות מבחני מובהקות (למשל, p-value) | קובע את רמת הביטחון הנדרשת להכרזה על שינוי מצב. סף נמוך מדי -> אזעקות שווא; סף גבוה מדי -> פספוס שינויים. |
| תדירות עדכון חישוב MMD | כל נר חדש של 30 דקות | מאזן בין עדכניות המידע לעלות החישובית. |

טבלה זו חיונית לתיעוד החלטות התכנון של רכיב מורכב וקריטי זה, ומסייעת בשחזור, ניפוי שגיאות וכיול עתידי של מנוע זיהוי מצבי השוק.

## **V. הפחתת ממדיות וייצוג וקטורי באמצעות רשתות נוירונים**

חלק זה מתעמק באופן שבו מטריצת הקלט הגבוה-ממדית וההטרוגנית עוברת טרנספורמציה לייצוג וקטורי דחוס, המותאם לקלט של סוכני למידת החיזוק.

### **A. רציונל ומטרות להטמעת מטריצת הקלט**

* **הבעיה:** מטריצת הקלט NxF, כפי שתוארה, עלולה להיות בעלת ממדיות גבוהה ולהכיל סוגי נתונים הטרוגניים (מאפיינים רציפים, קטגוריאליים, מבוססי-אירועים עם סקלות והתפלגויות שונות).1
* **מטרה 1: הפחתת ממדיות:** יצירת ייצוג וקטורי רציף, דחוס ובעל ממדיות נמוכה יותר.1
* **מטרה 2: הפחתת הטרוגניות:** מיפוי מאפיינים מגוונים למרחב משותף ואחיד יותר, שבו ניתן ללמוד יחסים ביניהם בקלות רבה יותר.1
* **מטרה 3: למידת מאפיינים (Feature Learning):** למידה אוטומטית של צירופים רלוונטיים ואינטראקציות לא-ליניאריות בין מאפייני הקלט הגולמיים.27
* **תועלת ללמידת חיזוק (RL):** אספקת ייצוג מצב יציב יותר, אינפורמטיבי ובעל ממדיות נמוכה יותר עבור סוכני ה-RL, מה שעשוי להאיץ את הלמידה ולשפר את איכות המדיניות הנלמדת.1
* רשת ההטמעה פועלת כממשק חיוני בין מרחב המאפיינים המהונדס והמורכב של מטריצת הקלט לבין רשתות המדיניות/ערך של סוכן ה-RL. יעילותה בדחיסת מידע והדגשת דפוסים רלוונטיים היא בעלת חשיבות עליונה.
  1. מטריצת הקלט NxF היא פוטנציאלית גדולה ומגוונת.
  2. סוכני RL עלולים להתקשות עם מרחבי מצב גבוהי-ממדיות והטרוגניים מאוד ("קללת הממדיות").
  3. רשת נוירונים להטמעה יכולה להקרין זאת לווקטור קטן ודחוס.
  4. וקטור זה אמור ללכוד את המידע הבולט ביותר למשימת המסחר.
  5. הדבר הופך את בעיית הלמידה של סוכן ה-RL לקלה יותר לניהול.

### **B. ניתוח השוואתי של ארכיטקטורות רשת נוירונים מועמדות**

בחירת הארכיטקטורה המתאימה לשכבת ההטמעה תלויה במאפייני הנתונים ובמטרות הלמידה.

* **רשתות קונבולוציה חד-ממדיות (1D-CNNs):**
  + **יתרונות:** יעילות בזיהוי דפוסים מקומיים ומוטיבים ברצפים.10 יכולות ללמוד מאפיינים היררכיים. יעילות חישובית יחסית.
  + **חסרונות:** שדה קליטה (receptive field) מוגבל, עשויות להתקשות בלכידת תלויות ארוכות-טווח אלא אם משתמשים בקונבולוציות מורחבות (dilated convolutions).30
  + **יישום:** יכולות לסרוק את מטריצת ה-NxF לאורך ציר הזמן (N) כדי לחלץ דפוסים זמניים מ-F המאפיינים, או על פני המאפיינים אם קיים סידור "מרחבי" משמעותי. מקורות 10 מציינים במפורש שימוש ב-1D CNNs להטמעת "טוקנים" (מאפיינים) ברשתות טרנספורמר עבור סדרות זמן פיננסיות, תוך למידת מידע זמני לכל נייר ערך/מאפיין.
* **רשתות רקורנטיות (LSTMs/GRUs):**
  + **יתרונות:** מתוכננות במיוחד עבור נתונים סדרתיים, יכולות ללכוד תלויות זמניות ארוכות-טווח והקשר.30 יכולות להתמודד עם רצפים באורך משתנה (אם כי כאן N קבוע).
  + **חסרונות:** עלולות להיות איטיות לאימון בשל עיבוד סדרתי, בעיות פוטנציאליות של התפוגגות/התפוצצות גרדיאנטים (אם כי מופחתות על ידי ארכיטקטורת LSTM/GRU).30 העיבוד הסדרתי מגביל מקבול.34
  + **יישום:** עיבוד רצף N וקטורי המאפיינים כדי לייצר מצב חבוי סופי או רצף של מצבים חבויים כהטמעה.
* **רשתות טרנספורמר (Transformers):**
  + **יתרונות:** משתמשות במנגנוני קשב (attention) ללכידת תלויות גלובליות ברצפים, ניתנות למקבול גבוה, ונחשבות למצב-האמנות (state-of-the-art) במשימות רצף רבות.3 גרסאות כמו ProbSparse Attention יכולות להתמודד ביעילות עם רצפים ארוכים.10
  + **חסרונות:** דורשות כמויות גדולות של נתונים לאימון, עלולות להיות יקרות חישובית עבור רצפים ארוכים מאוד (אם כי קיימות גרסאות יעילות), ועשויות לדרוש כיול קפדני.3
  + **יישום:** התייחסות לרצף N וקטורי המאפיינים כאל "טוקנים" קלט, מה שמאפשר למנגנון הקשב העצמי לשקלל את חשיבותם של צעדי זמן ומאפיינים שונים.
* **אוטואנקודרים (AE) / אוטואנקודרים וריאציוניים (VAE):**
  + **יתרונות:** למידה בלתי מונחית של ייצוגים דחוסים. VAEs מספקים מרחב סמוי (latent space) הסתברותי.28 יכולים לשמש לחילוץ מאפיינים והפחתת ממדיות. אוטואנקודרים גנרטיביים במסגרת DERL 28 לומדים הטמעות הלוכדות דינמיקות מעבר.
  + **חסרונות:** אוטואנקודרים סטנדרטיים עשויים שלא לעבור אופטימיזציה מפורשת עבור יכולת חיזוי ספציפית למשימה, אלא אם כן הם חלק ממערכת מקצה-לקצה. VAEs יכולים להיות מורכבים לאימון.
  + **יישום:** אימון AE/VAE על מטריצות הקלט NxF כדי ללמוד הטמעה קומפקטית משכבת צוואר הבקבוק. אם מאומנים מקצה-לקצה עם RL (כמו ב-DERL), הם לומדים הטמעות רלוונטיות למשימה.
* **שילוב ארכיטקטורות:** מודלים היברידיים (למשל, CNN-LSTM 10, LSTM-GNN 31, VAE-Transformer-LSTM 38) נפוצים כדי למנף את החוזקות של כל רכיב. לדוגמה, שכבות CNN לחילוץ מאפיינים ראשוני, ואחריהן LSTM/Transformer למודליזציה של הרצף.
* בחירת ארכיטקטורת ההטמעה אינה בלתי תלויה באופי מטריצת הקלט (NxF). אם N גדול מאוד (רצפים ארוכים), טרנספורמרים עם קשב יעיל או LSTMs עשויים להיות מועדפים. אם דפוסים מקומיים בזמן חשובים מאוד, CNNs חזקים. ההטרוגניות של F המאפיינים משנה גם היא; ארכיטקטורות מסוימות עשויות להתמודד טוב יותר עם סוגי מאפיינים מגוונים לאחר נרמול ראשוני.
  1. יש להטמיע את מטריצת ה-NxF.
  2. לארכיטקטורות רשת נוירונים שונות יש חוזקות שונות עבור נתונים סדרתיים.
  3. CNNs מצטיינים בחילוץ מאפיינים מקומיים.
  4. LSTMs מצטיינים בתלויות זמניות ארוכות-טווח.
  5. טרנספורמרים מצטיינים בתלויות גלובליות באמצעות קשב וניתנים למקבול טוב.
  6. אוטואנקודרים לומדים ייצוגים דחוסים, ולוכדים פוטנציאלית מבנה בסיסי.
  7. הבחירה "הטובה ביותר" תלויה במאפיינים הספציפיים של מטריצת ה-NxF (למשל, אורך הרצף N, אופי המאפיינים F, חשיבות דפוסים מקומיים לעומת גלובליים) ובמשימת ה-RL במורד הזרם.
  8. מודלים היברידיים מציעים לעיתים קרובות את הטוב מכל העולמות.
* שאיפתו של המשתמש לאמן את רשת ההטמעה באמצעות שגיאת TD מסוכן ה-RL (אימון מקצה-לקצה) משמעה שרשת ההטמעה אינה רק מעבד-מקדים קבוע. היא לומדת באופן פעיל ליצור ייצוגים שהם *שימושיים* לקבלת ההחלטות של סוכן ה-RL ולמקסום התגמול.1 הדבר הופך ארכיטקטורות כמו אלו שב-DERL 28 או DeepTrader 45 לרלוונטיות במיוחד.
  1. שאילתת משתמש: "האימון של הרשת הוא מה-TD error לצורך העניין של הסוכנים".
  2. הדבר מרמז על אימון מקצה-לקצה שבו אות הלמידה של סוכן ה-RL (שגיאת TD) מתפשט לאחור דרך רשתות המדיניות/ערך שלו *וגם* דרך רשת ההטמעה.
  3. משמעות הדבר היא שרשת ההטמעה אינה מאומנת רק על משימת שחזור או למידה עצמית-מונחית נפרדת.
  4. היא לומדת לייצר הטמעות המסייעות ישירות לסוכן ה-RL למקסם תגמולים.
  5. זהו רעיון רב עוצמה, כפי שנראה במודלים כמו DERL 28, שבהם האוטואנקודר הוא חלק מלולאת ה-RL ומסתגל.

### **טבלה 4: השוואת ארכיטקטורות רשת נוירונים להטמעת מטריצת קלט**

| **ארכיטקטורה** | **חוזק עיקרי להטמעת סדרות זמן פיננסיות** | **חולשות/אתגרים עיקריים** | **התאמה למטריצת NxF (מורכבות זמן ומאפיינים)** | **הערות על אימון מקצה-לקצה עם RL** |
| --- | --- | --- | --- | --- |
| 1D-CNN | חילוץ מאפיינים מקומיים ודפוסים זמניים קצרים, יעילות חישובית. | שדה קליטה מוגבל לתלויות ארוכות טווח (אלא אם כן משתמשים בקונבולוציות מורחבות). | טובה לרצפים קצרים-בינוניים (N), יכולה לטפל במספר מאפיינים (F) במקביל. | יכולה לשמש כשכבת חילוץ מאפיינים ראשונית לפני רשת RL עמוקה יותר; הגרדיאנטים יכולים לזרום דרכה. |
| LSTM/GRU | לכידת תלויות זמניות ארוכות-טווח והקשר ברצפים. | עיבוד סדרתי מגביל מקבול, עלולה להיות איטית לאימון, רגישות להיפר-פרמטרים. | מתאימה היטב לרצפים ארוכים (N), יכולה לעבד וקטורי מאפיינים (F) בכל צעד זמן. | סטנדרטית ב-RL מבוסס-רצף; המצב החבוי האחרון משמש לעיתים קרובות כהטמעה. |
| Transformer | לכידת תלויות גלובליות באמצעות קשב עצמי, מקבול גבוה. | דורשת נתונים רבים, יקרה חישובית לרצפים ארוכים מאוד (ללא גרסאות יעילות), כיול מורכב. | יכולה להתמודד עם רצפים ארוכים (N) ומספר רב של מאפיינים (F), במיוחד עם מנגנוני קשב יעילים. | פחות נפוצה כשכבת הטמעה ישירה ב-RL קלאסי בהשוואה ל-LSTM, אך אפשרית; הקשב יכול ללמוד אילו חלקים מההיסטוריה (N) והמאפיינים (F) רלוונטיים. |
| Autoencoder/VAE | למידה בלתי מונחית של ייצוגים דחוסים, הפחתת רעשים. VAEs לוכדים התפלגות סמויה. | הטמעה סטנדרטית אינה מותאמת למשימת RL ספציפית אלא אם מאומנת מקצה-לקצה. | יכולה לדחוס את כל מטריצת NxF או כל וקטור מאפיין F בנפרד. | אימון מקצה-לקצה (כמו ב-DERL 28) הופך את ההטמעה לרלוונטית למשימה. |
| היברידי (למשל, CNN+LSTM) | שילוב חוזקות (למשל, CNN לחילוץ מאפיינים מקומיים, LSTM למודליזציה של תלויות זמניות). | מורכבות מוגברת בתכנון ובאימון. | יכולה להתאים למגוון רחב של מאפייני NxF על ידי התאמת הרכיבים. | נפוץ; מאפשר לכל רכיב להתמחות במה שהוא טוב בו, תוך שהגרדיאנטים זורמים דרך כל המבנה. |

טבלה זו מספקת השוואה מובנית כדי לסייע למשתמש לבחור את ארכיטקטורת הרשת הנוירונית המתאימה ביותר לשכבת ההטמעה שלו. היא מקשרת בין בחירות ארכיטקטוניות לצרכים הספציפיים של נתוני סדרות זמן פיננסיות ולמטרה של אימון RL מקצה-לקצה.

## **VI. סוכני למידת חיזוק לביצוע עסקאות וניהול סיכונים**

חלק זה מפרט את מערכת ה-MARL אשר צורכת את הייצוג הווקטורי המוטמע לצורך קבלת החלטות מסחר.

### **A. ארכיטקטורה (למשל, CTDE עם מדיניות משותפת) וייצוג מצב הסוכן**

* **העדפת המשתמש:** מערכת MARL שבה סוכנים חולקים מדיניות משותפת אך עשויים לקבל "תצוגות" (views) שונות של הנתונים המוטמעים.1 הדבר עולה בקנה אחד עם הרעיון של חלוקת נתונים דינמית המבוססת על מצב שוק.1
* **CTDE (Centralized Training with Decentralized Execution):** פרדיגמה מתאימה. במהלך האימון, מבקר (critic) מרכזי יכול להשתמש במידע גלובלי (או תצפיות/פעולות מצטברות של סוכנים) כדי לאמן את המדיניות המשותפת. במהלך הביצוע, כל סוכן משתמש בתצפית המקומית (המוטמעת) שלו ובמדיניות המשותפת.1
* **מצב הסוכן:** הווקטור הפלט משכבת ההטמעה של הרשת הנוירונית משמש כייצוג המצב st​ עבור סוכן/סוכני ה-RL.1
* **ספריות:** Ray RLlib ו-PettingZoo מתאימות למימוש מערכות MARL כאלו.1 MAPPO הוא אלגוריתם רלוונטי עבור CTDE עם מדיניות משותפת.1
* ה"תצוגות השונות" עבור סוכנים החולקים מדיניות, אם רשת ההטמעה עצמה משותפת, מרמזות על כך שה*קלט* לרשת ההטמעה המשותפת עשוי להיות שונה לכל סוכן (כלומר, מטריצות Mi​ שונות לפני ההטמעה), או שסוכנים מפנים קשב באופן סלקטיבי לחלקים שונים של וקטור מוטמע משותף גלובלית. הבהרת המשתמש שהרשת הנוירונית *אינה* מיועדת לחלוקת נתונים אלא להטמעת הנתונים *שכבר חולקו/ספציפיים-לתצוגה* תומכת באפשרות הראשונה.1
  1. המשתמש מעוניין ב-MARL עם מדיניות משותפת אך סוכנים עשויים לקבל "תצוגות" שונות.
  2. המשתמש גם הבהיר שרשת הנוירונים מיועדת להטמעה, לא לחלוקת נתונים.
  3. הדבר מרמז שחלוקת הנתונים ל-Mi​ (פוטנציאלית דינמית ומבוססת-מצב שוק) מתרחשת *לפני* שלב ההטמעה.
  4. כל סוכן i מעבד את ה-Mi​ שלו דרך רשת נוירונים להטמעה (שעשויה להיות בעלת משקולות משותפות לכל הסוכנים אם הם הומוגניים ביכולתם) כדי לקבל את המצב si​ שלו.
  5. כל הסוכנים משתמשים אז באותה מדיניות π(ai​∣si​).
  6. "שיתוף פעולה מושלם" נוגע אז לתוצאה הגלובלית של פעולות נגזרות-מקומיות אלו מלוגיקה משותפת.

### **B. עיצוב מרחב הפעולות ולמידת מדיניות מתשומות אסטרטגיות**

* **מרחב הפעולות:** חייב לאפשר לסוכן לבצע עסקאות. סביר להניח שיידרש מרחב פעולות היברידי:
  + **בדיד:** כיוון עסקה (לונג, שורט, החזק/סגור). סוג פקודה (שוק, גבול – אף שהאסטרטגיה של המשתמש אינה מציינת פקודות גבול, זהו שיקול נפוץ ב-RL).
  + **רציף/בדיד-מרובה-רמות:** גודל פוזיציה. רמות עצירת הפסד (stop-loss), רמות לקיחת רווח (take-profit) (אם נקבעות דינמית על ידי הסוכן).
* **למידת מדיניות:** סוכן ה-RL לומד מדיניות π(action∣state) הממפה את וקטור המצב המוטמע לפעולה אופטימלית. המדיניות אמורה ללמוד באופן משתמע לפרש את הרכיבים האסטרטגיים המקודדים במצב (אותות אינדיקטורים, עוצמת LVN, גיאומטריית משולש הרמוני, מצב שוק).
* **ניהול סיכונים במדיניות:** הסוכן צריך ללמוד לא רק *מתי* לסחור אלא גם *כמה* לסחור, תוך שילוב ניהול סיכונים.
* רשת המדיניות צריכה להיות מורכבת מספיק כדי לפענח את המידע העשיר מהווקטור המוטמע, כולל הרמזים הגיאומטריים העדינים מהמשולשים ההרמוניים ועוצמת ה-LVN ההקשרית, ולמפות זאת לפעולות מתאימות, במיוחד לגודל הפוזיציה.
  1. מצב הסוכן הוא הווקטור המוטמע המכיל את כל המידע האסטרטגי.
  2. פעולת הסוכן כוללת את גודל הפוזיציה.
  3. המשתמש מעוניין שגודל הפוזיציה ישתנה בהתבסס על עוצמת LVN ואיכות הצורה ההרמונית.
  4. רשת המדיניות חייבת ללמוד מיפוי מורכב ולא-ליניארי זה מהמאפיינים במצב המייצגים LVN/גיאומטריה לרכיב גודל הפוזיציה של הפעולה.
  5. זוהי משימת למידה מאתגרת המטמיעה ישירות ניהול סיכונים במדיניות.

### **C. ניהול סיכונים דינמי באמצעות קביעת גודל פוזיציה (עוצמת LVN, איכות הרמונית)**

* **דרישת המשתמש:** גודל הפוזיציה צריך להיות מושפע מ"עוצמת תחנת" LVN ו"איכות הצורה ההרמונית".1
* **מימוש:** מאפיינים (מכומתים) אלו הם חלק ממצב הסוכן. רשת המדיניות של סוכן ה-RL תלמד להפיק גודל פוזיציה (או פרמטר הקובע את גודל הפוזיציה) כחלק מפעולתה, המותנה במאפיינים אלו.
* **פונקציית תגמול:** פונקציית התגמול חייבת להיות מתוכננת כך שתנחה באופן משתמע את הסוכן ללמוד התנהגות זו. לדוגמה, רווחים גדולים יותר מפוזיציות בגודל מתאים (המבוססות על אותות חזקים) אמורים להיות מתוגמלים, בעוד שהפסדים גדולים יותר מפוזיציות בגודל לא מתאים (במיוחד על אותות חלשים) אמורים להיענש בחומרה רבה יותר מאשר אם היה נעשה שימוש בגודל קטן יותר.
* זוהי צורה מתוחכמת של ניהול סיכונים משולב. הסוכן אינו פועל רק על פי חוקים קבועים לקביעת גודל פוזיציה, אלא לומד אסטרטגיית קביעת גודל אופטימלית המבוססת על מאפייני קלט דינמיים ומורכבים. זהו צעד משמעותי מעבר למערכות מסורתיות.
  1. המשתמש מעוניין בקביעת גודל פוזיציה דינמית בהתבסס על עוצמת LVN ואיכות הצורה ההרמונית.
  2. אלו הם מאפיינים במצב הסוכן.
  3. מרחב הפעולות של הסוכן כולל את גודל הפוזיציה.
  4. אלגוריתם ה-RL (למשל, PPO, SAC) יאמן את רשת המדיניות למפות מצבים לפעולות (כולל גודל פוזיציה) כדי למקסם תגמול מצטבר.
  5. אם פונקציית התגמול משקפת נכונה את הרצון לרווחים גדולים יותר עם סיכון מוצדק היטב ומענישה על הפסדים ממינוף-יתר על אותות חלשים, הסוכן אמור ללמוד את התנהגות קביעת הגודל הרצויה.

### **טבלה 5: תצורת מצב, פעולה ומדיניות של סוכן MARL**

| **רכיב** | **תיאור / פרטים** |
| --- | --- |
| ייצוג מצב (st​) | וקטור מוטמע מרשת ההטמעה הנוירונית, המקודד את כל המאפיינים ממטריצת הקלט Mi​ (הייקן אשי, אינדיקטורים, LVN, גיאומטריה הרמונית, מצב שוק MMD). |
| מרחב פעולות (at​) | היברידי: {סוג\_עסקה: [קנייה, מכירה, החזקה/סגירה], מקדם\_גודל\_פוזיציה: [רציף, למשל 0.1-2.0 כפול יחידת בסיס], (אופציונלי) יעד\_רווח\_יחסי: [רציף], (אופציונלי) עצירת\_הפסד\_יחסית: [רציף]} |
| ארכיטקטורת רשת מדיניות מוצעת | רשת Feedforward (MLP) על הווקטור המוטמע st​. "ראשים" (output heads) נפרדים עבור רכיבי הפעולה הבדידים (עם הפעלת softmax) והרציפים (עם הפעלה מתאימה, למשל tanh מותאם לטווח). |
| אלגוריתם RL לדוגמה | MAPPO (עבור CTDE עם מדיניות משותפת), PPO, SAC (עם התאמות למרחב פעולות היברידי). |
| היפר-פרמטרים מרכזיים ל-RL | קצב למידה (learning rate), מקדם היוון (gamma γ), גודל אצווה (batch size), אנטרופיה (לעידוד חקירה). |

טבלה זו מבהירה את הקלטים והפלטים הישירים של סוכן ה-RL ומספקת נקודת מוצא לבחירת אלגוריתמי RL וארכיטקטורות רשת מתאימות למדיניות. היא מקשרת את עיבוד הנתונים במעלה הזרם לקבלת ההחלטות במורד הזרם.

## **VII. צנרת האימון המשולבת: מנתונים להחלטת סוכן**

חלק זה מתווה את זרימת המידע והלמידה מקצה-לקצה בתוך המערכת.

### **A. זרימת נתונים וטרנספורמציה מקצה-לקצה**

1. **רכישת נתוני נכס גולמיים** (OHLCV של חוזים עתידיים).1
2. **המרת הנתונים לנרות הייקן אשי** של 30 דקות ו-5 דקות.1
3. **חישוב אינדיקטורים על נרות הייקן אשי:**
   * 30 דקות: MLMI, NW-RQK.1
   * 5 דקות: FVG.1
4. **ניתוח פרופיל שוק לזיהוי LVNs** (מנתוני 30 דקות) וכימות "עוצמת התחנה".1
5. **זיהוי מצב שוק באמצעות MMD** (מנתוני הייקן אשי של 30 דקות) לקבלת תווית/ציון מצב.1
6. **זיהוי התיישרות אותות (סינרגיות) ויצירת משולשים הרמוניים** על גרף 5 דקות; מיצוי מאפיינים גיאומטריים.1
7. **בניית מטריצת הקלט NxF** עבור כל צעד זמן, המשלבת את כל המאפיינים משלבים 2-6.1 כאן עשויה להתרחש חלוקה דינמית המבוססת על מצב השוק ליצירת "תצוגות" ספציפיות לסוכן Mi​.1
8. **הזנת מטריצת הקלט (Mi​) לשכבת ההטמעה של הרשת הנוירונית.**
9. **רשת ההטמעה הנוירונית מפיקה ייצוג וקטורי דחוס** (מצב הסוכן st​).1
10. **רשת המדיניות של סוכן ה-RL מקבלת את st​ ומפיקה פעולה at​** (כולל סוג עסקה, גודל פוזיציה).
11. **הפעולה at​ מבוצעת בסביבת השוק** (מדומה או חיה).
12. **הסביבה מחזירה תגמול rt​ ואת המצב הבא st+1​** (דרך המאפיינים הגולמיים שלו, ואז עיבוד מחדש משלב 1/2).

* זרימה מפורטת זו מדגישה את שלבי הטרנספורמציה וההפשטה הרבים של המאפיינים. הייצוג הווקטורי הסופי st​ הוא מעובד מאוד, וסוכן ה-RL לומד על בסיס מצב מופשט זה. איכות כל שלב קודם משפיעה ישירות על איכות st​.
  1. המערכת מורכבת משלבים רבים.
  2. הדמיית זרימת הנתונים מסייעת להבין תלויות.
  3. נתונים גולמיים הופכים לנתוני הייקן אשי.
  4. נתוני הייקן אשי מזינים אינדיקטורים ו-MMD.
  5. אותות יוצרים משולשים.
  6. כל אלו הופכים למאפיינים במטריצה.
  7. המטריצה מוטמעת.
  8. הווקטור המוטמע הוא מצב הסוכן.
  9. הסוכן פועל, מקבל תגמול.
  10. מחזור זה מניע למידה.
  11. כל פגם בשלב מוקדם (למשל, חישוב שגוי של הייקן אשי, MMD מכויל גרוע) יפיץ רעש או שגיאות למצב הסוכן, ויפגע בלמידה.

### **B. לולאת למידת הסוכן ואימון רשת מבוסס שגיאת TD**

* **שאלת הליבה של המשתמש:** "האימון של הרשת הוא מהTD error לצורך העניין של הסוכנים ובעצם אנחנו מהנדסים פה רוטינת אימון ושפה משותפת של כל הפרויקט" (שאילתת משתמש1).
* **אלגוריתם RL:** ייעשה שימוש באלגוריתם כגון PPO או SAC. אלגוריתמים אלו מחשבים פונקציית הפסד המבוססת על ניסיון הסוכן (מצבים, פעולות, תגמולים). עבור שיטות שחקן-מבקר (actor-critic), הדבר כולל הערכת פונקציות ערך (Q-value או state-value V) וגרדיאנטי מדיניות.
* **שגיאת TD (Temporal Difference Error):** מושג ליבה, המייצג את ההפרש בין הערך המוערך של מצב (או זוג מצב-פעולה) לבין ערך מטרה מדויק יותר (הנגזר בדרך כלל מהתגמול הנצפה ומהערך המוערך של המצב הבא). למשל, δt​=rt​+γV(st+1​)−V(st​).42
* **התפשטות לאחור (Backpropagation):** ההפסד הנגזר משגיאות TD (או ממטרות גרדיאנט מדיניות) מתפשט לאחור דרך רשתות המדיניות והערך של סוכן ה-RL כדי לעדכן את משקולותיהן.42
* **אימון מקצה-לקצה של רשת ההטמעה:** באופן קריטי, אם רשת ההטמעה היא חלק מהארכיטקטורה הכוללת של רשת הסוכן (כלומר, הפלט שלה מוזן ישירות לרשתות המדיניות/ערך), אזי הגרדיאנטים מהפסד ה-RL יתפשטו לאחור *גם דרך* רשת ההטמעה.1
  + משמעות הדבר היא שרשת ההטמעה לומדת לייצר ייצוגים שאינם רק טובים לשחזור קלט (כמו אוטואנקודר סטנדרטי) אלא שימושיים באופן ספציפי עבור סוכן ה-RL למקסום התגמול המצטבר שלו.
  + מסגרת DERL 28 ו-DeepTrader 45 הן דוגמאות לאימון מקצה-לקצה שבו מחלצי/מטמיעי מאפיינים מאומנים יחד עם מדיניות ה-RL.
* אימון מקצה-לקצה זה הוא רב עוצמה מכיוון שהוא מיישר את מטרות למידת הייצוג (הטמעה) עם מטרות קבלת ההחלטות (RL). "השפה המשותפת" שהמשתמש מתייחס אליה היא תהליך אופטימיזציה משותף זה שבו ההטמעה מסתגלת לשרת את צרכי הסוכן.
  1. סוכן RL לומד על ידי מזעור פונקציית הפסד הנגזרת מתגמולים (ולעיתים קרובות כוללת שגיאות TD).
  2. הפסד זה מתפשט לאחור כדי לעדכן את משקולות הרשת.
  3. אם פלט רשת ההטמעה הוא הקלט לרשתות סוכן ה-RL, הן יוצרות גרף חישוב אחד גדול.
  4. גרדיאנטים יזרמו מהפסד ה-RL כל הדרך חזרה למשקולות רשת ההטמעה.
  5. לפיכך, רשת ההטמעה לומדת ליצור מאפיינים המשפרים את ביצועי סוכן ה-RL במשימת המסחר, במקום לעבור אופטימיזציה למטרה בלתי תלויה.
* **שאלת המשתמש לגבי נתוני אימון:** "האם אני מכניס דאטה נקייה כמו נתונים של אינדיקטורים נכס והכל נקי בלי חישוב והסוכנים רצים על הפוליסה... או שמע אני מכניס נתונים של אחרי בקטסט והסוכנים לומדים".
  + **הבהרה:** סוכנים מאומנים בדרך כלל על נתונים היסטוריים (שאמורים להיות נקיים) המדמים את סביבת השוק. ה"מדיניות" היא מה שהסוכן *לומד* במהלך אימון זה. בדיקה לאחור (backtest) היא דרך *להעריך* מדיניות נלמדת על נתונים היסטוריים שהיא לא ראתה במהלך האימון. לכן, סוכנים לומדים מנתונים היסטוריים "נקיים" על ידי אינטראקציה עם סביבה מדומה. הם אינם לומדים "אחרי בדיקה לאחור" במובן של שימוש בתוצאות הבדיקה לאחור כקלט ישיר, אלא שתהליך הבדיקה לאחור (במיוחד אופטימיזציית walk-forward) כולל אימון מחדש איטרטיבי על נתונים היסטוריים חדשים.
* ההבחנה בין אימון על "נתונים נקיים" לבין "נתונים אחרי בדיקה לאחור" דורשת הבהרה. סוכני RL לומדים על ידי אינטראקציה עם סביבה המספקת מצבים, תגמולים ומצבים הבאים. סביבה זו בנויה בדרך כלל באמצעות נתוני שוק היסטוריים (שאמורים להיות נקיים ומעובדים לפורמט מטריצת NxF). ה"מדיניות" היא פלט תהליך הלמידה. בדיקה לאחור מיועדת להערכת מדיניות נלמדת זו. ייתכן שהמשתמש חושב האם הסוכן לומד מאותות גולמיים או מתוצאות רווח והפסד של אסטרטגיה מוגדרת מראש; ב-RL, הוא לומד על ידי נקיטת פעולות והתבוננות בתגמולים המבוססים על פעולות אלו, ובכך מגלה למעשה אסטרטגיה משלו למקסום התגמול.

### **טבלה 6: סקירת לולאת האימון המשולבת**

| **שלב** | **תיאור התהליך** | **קלטים מרכזיים** | **פלטים מרכזיים** | **רכיבים/ספריות רלוונטיים** |
| --- | --- | --- | --- | --- |
| 1 | קליטת נתונים גולמיים והמרתם להייקן אשי | נתוני OHLCV גולמיים | נרות הייקן אשי (30 דק', 5 דק') | צנרת עיבוד נתונים (מותאמת אישית) |
| 2 | חישוב אינדיקטורים, LVN, MMD, וגיאומטריה הרמונית | נרות הייקן אשי, נתוני פרופיל שוק | ערכי אינדיקטורים, מאפייני LVN, ציון/תווית מצב MMD, מאפיינים גיאומטריים של משולש | ספריות אנליזה טכנית (למשל, TA-Lib, pandas-ta), market\_profile, issaz/signature-regime-detection, לוגיקה מותאמת אישית |
| 3 | בניית מטריצת קלט Mi​ (פוטנציאל דינמי) | כלל המאפיינים משלב 2 | מטריצת NxF עבור כל סוכן/תצוגה | לוגיקת הרכבת מאפיינים |
| 4 | הטמעת רשת נוירונים (st​=Emb(Mi​)) | מטריצת קלט Mi​ | וקטור מצב מוטמע st​ | רשת הטמעה (למשל, CNN, LSTM, Transformer) |
| 5 | פעולת סוכן (at​=π(st​)) | מצב סוכן st​ | פעולה at​ (כיוון, גודל פוזיציה וכו') | רשת מדיניות RL (בתוך RLlib) |
| 6 | צעד סביבה ותגמול (rt​,st+1​) | פעולה at​, מצב שוק נוכחי | תגמול rt​, מצב גולמי הבא (שיוזן חזרה לשלב 1/2) | סביבת מסחר מדומה (PettingZoo) |
| 7 | חישוב שגיאת TD | rt​,V(st​),V(st+1​) (או Q-values) | δt​ | אלגוריתם RL (בתוך RLlib) |
| 8 | התפשטות לאחור ועדכון משקולות | δt​, גרדיאנטים | משקולות מעודכנות לרשת ההטמעה ולרשתות המדיניות/ערך | מנוע אופטימיזציה (למשל, Adam ב-PyTorch/TensorFlow דרך RLlib) |

טבלה זו ממפה באופן חזותי ותיאורי את כל מחזור האימון המורכב, ומבהירה כיצד רכיבים שונים מקיימים אינטראקציה וכיצד אות הלמידה (שגיאת TD) מתפשט כדי לאמן הן את רשת ההטמעה והן את מדיניות סוכן ה-RL. הדבר עונה ישירות על חלק מרכזי בשאילתת המשתמש.

## **VIII. רובסטיות מערכת, בדיקות לאחור ושיקולים עתידיים**

חלק זה מתייחס להיבטים קריטיים בהבטחת אמינות המערכת לקראת מסחר חי.

### **A. אסטרטגיות למניעת התאמת יתר והבטחת יכולת הכללה**

* **אופטימיזציית Walk-Forward:** חיונית עבור נתוני סדרות זמן לבדיקת ביצועים מחוץ-לדגימה ולהסתגלות לדינמיקות שוק משתנות.1
* **רגולריזציה:** רגולריזציית L1/L2 עבור משקולות רשת נוירונים, שכבות dropout.1
* **אימות צולב (Cross-Validation):** שימוש בטכניקות אימות צולב המתאימות לסדרות זמן.
* **נתוני אימון מגוונים:** אימון על פני מגוון תנאי שוק (מצבים) ופוטנציאלית נכסים שונים (אם המודל מיועד להיות כללי).
* **עיצוב תגמול (Reward Shaping):** תכנון קפדני של פונקציית התגמול כדי להימנע מתגמול על מתאמים אקראיים או מדיניויות מורכבות מדי.46
* **עקרון הפשטות:** בעוד שהמערכת מורכבת, רכיבים בודדים צריכים להיות פשוטים ככל האפשר תוך השגת מטרתם.
* הערת המשתמש לגבי כך שהמדיניות "טובה" (שיעור הצלחה של 46% הוזכר) ורק זקוקה לשיפור בניהול סיכונים היא נקודת התחלה טובה, אך מערכת RL מלאה תלמד מחדש את המדיניות. המיקוד צריך להיות על למידה רובסטית ויכולת הכללה, ולא רק על התאמה לאותות "טובים" מהעבר.
  1. המשתמש מזכיר שיעור הצלחה של 46% עבור סינרגיות קיימות וצורך בשיפור ניהול סיכונים.
  2. סוכן RL ילמד מדיניות משלו למקסום תגמול, שעשויה להיות שונה מהפרשנות הקיימת של הסינרגיות.
  3. המפתח הוא שסוכן ה-RL ילמד מדיניות *רובסטית*.
  4. התאמת יתר היא סיכון מרכזי בלמידת מכונה פיננסית.1
  5. טכניקות כמו אופטימיזציית walk-forward, רגולריזציה ונתוני אימון מגוונים הן חיוניות ליכולת הכללה.

### **B. שיקולים לגבי סופרפוזיציה בהשראת קוונטים (בהתאם לעניין המשתמש)**

* **מטרת המשתמש:** שיפור רובסטיות ואמינות.1
* **אנלוגיות קלאסיות:** שיטות אנסמבל, מדיניויות הסתברותיות, התחייבות מאוחרת דמוית-QIDDM.1
* **שילוב פוטנציאלי:** יכול להיות מיושם ברמת הסוכן (למשל, סוכן המחזיק מספר השערות לגבי הפעולה הטובה ביותר) או ברמת המערכת (אנסמבל של סוכנים).
* **תזמון החלטה:** המשתמש הביע עניין לדחות חקירה מעמיקה של נושא זה עד להבנה טובה יותר של מערכת ה-MARL הבסיסית.1
* רעיונות סופרפוזיציה, אף שהם מופשטים, מציעים גישות אלגוריתמיות קלאסיות קונקרטיות (כמו אנסמבלים) הידועות בשיפור רובסטיות. התחייבות מאוחרת דמוית-QIDDM מעניינת במיוחד למסחר, שכן היא מפרמלת את הרעיון של המתנה לרמת ודאות גבוהה יותר.
  1. המשתמש מתעניין בסופרפוזיציה בהשראת קוונטים לשיפור רובסטיות.
  2. אין הכוונה לחישוב קוונטי אמיתי.
  3. קיימות אנלוגיות קלאסיות: אנסמבלים, פעולות הסתברותיות, התחייבות מאוחרת (QIDDM).
  4. שיטות אלו יכולות לשפר רובסטיות על ידי גיוון אסטרטגיות או דחיית החלטות עד לוודאות רבה יותר.
  5. נקודות המוצא המעשיות ביותר הן אנסמבלים או פלטים הסתברותיים מרשת המדיניות. QIDDM מתקדם יותר אך מתיישב היטב עם אינטואיציית המסחר.

## **IX. מסקנות והמלצות אסטרטגיות**

המחקר הנוכחי הציג תוכנית מקיפה ושאפתנית לפיתוח מערכת מסחר אלגוריתמית מרובת סוכנים (MARL) המיועדת למסחר בחוזים עתידיים. המערכת משלבת באופן חדשני אסטרטגיית מסחר ייחודית המבוססת על אינדיקטורים הפועלים על נרות הייקן אשי, ניתוח צמתי נפח נמוך (LVN) מפרופיל שוק, פרשנות גיאומטרית של צירופי אותות כ"צורות הרמוניות" (משולשים), זיהוי מצבי שוק באמצעות Higher-Rank MMD, ושכבת הטמעה מבוססת רשת נוירונים הנלמדת מקצה-לקצה כחלק אינטגרלי מסוכני ה-MARL.

**המלצות מרכזיות לפיתוח:**

1. **פיתוח מודולרי ואימות קפדני:**
   * יש לפתח ולבדוק כל רכיב מרכזי במערכת (צנרת נתונים והמרת הייקן אשי, חישוב כל אינדיקטור בנפרד, מנוע זיהוי LVN וכימות "עוצמת תחנה", מנוע MMD לזיהוי מצבי שוק, מנגנון חילוץ מאפיינים גיאומטריים ממשולשים הרמוניים, שכבת ההטמעה הנוירונית, ומסגרת סוכני ה-MARL) באופן עצמאי ככל הניתן לפני שילובו במערכת הכוללת.
   * יש להקדיש תשומת לב מיוחדת לאימות הסטטיסטי של כוח הניבוי של המאפיינים הגיאומטריים הנגזרים מהמשולשים ההרמוניים, כדי למנוע התאמת יתר לדפוסים אקראיים.
   * יש להבהיר ולהגדיר במדויק האם חישובי FVG ו-LVN יבוצעו על נתוני הייקן אשי (ויצריכו התאמה של ההגדרות הסטנדרטיות) או על נתונים גולמיים (ויצריכו מנגנון יישור עם אותות הייקן אשי).
2. **תכנון קפדני של מטריצת הקלט (NxF) ושכבת ההטמעה:**
   * יש להגדיר בבירור את כל F המאפיינים שייכללו בווקטור הקלט בכל צעד זמן (5 דקות), כולל אופן הייצוג של מאפיינים מבוססי-אירועים (כגון משולשים הרמוניים) ומאפיינים המתעדכנים בתדירות נמוכה יותר (כגון פלט MMD או מאפייני LVN).
   * יש לבחור בקפידה את אורך הרצף N, תוך איזון בין הצורך בהקשר היסטורי מספק לבין מגבלות חישוביות ויכולת הלמידה של הרשת.
   * יש לבצע נרמול וסקיילינג קפדניים של כל המאפיינים במטריצת הקלט.
   * יש לבחון ולהשוות ארכיטקטורות רשת נוירונים שונות (1D-CNN, LSTM, Transformer, Autoencoder, או שילובים היברידיים) עבור שכבת ההטמעה, בהתחשב באופי הנתונים ובמטרה של אימון מקצה-לקצה עם אות שגיאת ה-TD של סוכני ה-RL.
3. **מיקוד בלמידה רובסטית של סוכני ה-MARL:**
   * יש לתכנן את פונקציית התגמול של הסוכנים כך שתעודד לא רק מקסום רווחים אלא גם ניהול סיכונים מושכל (למשל, על ידי שילוב מדדים מותאמי-סיכון או עונשים על משיכות גדולות), ובפרט את הלמידה של קביעת גודל פוזיציה דינמי בהתבסס על "עוצמת תחנת" LVN ו"איכות הצורה ההרמונית".
   * יש ליישם טכניקות למניעת התאמת יתר, כגון רגולריזציה של רשתות הסוכנים ושימוש באופטימיזציית Walk-Forward קפדנית במהלך הבדיקות לאחור.
   * יש להבטיח שהמדיניות המשותפת של הסוכנים, אף שהיא מיושמת על "תצוגות" נתונים שונות (הנובעות מחלוקה דינמית מבוססת-מצב שוק), מובילה לפעולות קוהרנטיות ברמה הגלובלית.
4. **בדיקות לאחור מקיפות ואיטרטיביות:**
   * יש לבצע בדיקות לאחור מקיפות תוך שימוש בטכניקת Walk-Forward Analysis כדי להעריך את ביצועי המערכת בתנאים משתנים ולהפחית את הסיכון להתאמת יתר.
   * יש לכלול מידול ריאליסטי של עלויות עסקה (עמלות, החלקה).
   * יש להעריך את ביצועי המערכת על פני מגוון רחב של מדדים (ולא רק רווח והפסד), כולל מדדים מותאמי-סיכון, אחוזי הצלחה, וביצועים ספציפיים לסינרגיות השונות ולמצבי שוק שונים.

**כיווני מחקר עתידיים:**

לאחר ביסוס מערכת הליבה, ניתן יהיה לחקור מספר כיוונים מתקדמים:

* **חלוקת נתונים דינמית ולמידה:** פיתוח מנגנונים שבהם חלוקת הנתונים ל"תצוגות" של הסוכנים אינה רק מבוססת-מצב שוק קבוע, אלא נלמדת או מותאמת באופן דינמי יותר על ידי המערכת עצמה.
* **אבולוציה של סינרגיות:** בחינת האפשרות שהמערכת תלמד או תפתח באופן אוטונומי סינרגיות חדשות או וריאציות על הסינרגיות הקיימות, מעבר לארבע שהוגדרו מראש.
* **שילוב עמוק יותר של עקרונות בהשראת סופרפוזיציה:** לאחר הערכת ביצועי מערכת ה-MARL הבסיסית, ניתן לחקור באופן מעמיק יותר את שילובן של טכניקות סופרפוזיציה קלאסיות (כגון אנסמבלים מתוחכמים או מנגנוני התחייבות מאוחרת דמויי-QIDDM) לשיפור נוסף של עמידות המערכת.
* **פרשנות ושקיפות:** פיתוח כלים וטכניקות להבנת תהליך קבלת ההחלטות של מערכת MARL מורכבת זו, במטרה להגביר את האמון בה ולהקל על איתור וטיפול בכשלים.

לסיכום, המערכת המוצעת מייצגת חזון שאפתני וחדשני למסחר אלגוריתמי. מימושה ידרוש מאמץ מחקרי ופיתוחי משמעותי, אך הפוטנציאל לקידום התחום ולהשגת ביצועי מסחר משופרים מצדיק את המאמץ. התקדמות הדרגתית, תוך התבססות על מחקר קיים וכלים מתקדמים, והתמודדות שיטתית עם האתגרים, היא המפתח להפיכת חזון זה למציאות.

#### Works cited

1. שילוב סופרפוזיציה במסחר אלגוריתמי
2. CMES | Free Full-Text | Deep Learning for Financial Time Series Prediction: A State-of-the-Art Review of Standalone and Hybrid Models, accessed on June 4, 2025, <https://www.techscience.com/CMES/v139n1/55114/html>
3. A Comprehensive Survey of Deep Learning for Time Series Forecasting: Architectural Diversity and Open Challenges - arXiv, accessed on June 4, 2025, <https://arxiv.org/pdf/2411.05793>
4. Non-parametric online market regime detection and regime clustering for multidimensional and path-dependent data structures - IDEAS/RePEc, accessed on June 4, 2025, <https://ideas.repec.org/p/arx/papers/2306.15835.html>
5. OMI Research Newsletter – August 2023 - Oxford Man Institute of Quantitative Finance, accessed on June 4, 2025, <https://oxford-man.ox.ac.uk/wp-content/uploads/2023/08/OMINewsLetter-Aug23.pdf>
6. koenecke.infosci.cornell.edu, accessed on June 4, 2025, <https://koenecke.infosci.cornell.edu/files/Deep_Learning_for_Time_Series_Tutorial.pdf>
7. Course:CPSC522/Financial Forecasting using LSTM Networks ..., accessed on June 4, 2025, <https://wiki.ubc.ca/Course:CPSC522/Financial_Forecasting_using_LSTM_Networks>
8. The Impact of Data Preprocessing and Attribute Scaling on Neural Network Accuracy for Loan Default Prediction - ResearchGate, accessed on June 4, 2025, <https://www.researchgate.net/publication/388530244_The_Impact_of_Data_Preprocessing_and_Attribute_Scaling_on_Neural_Network_Accuracy_for_Loan_Default_Prediction>
9. 10 Stats: Effective Min-Max Scaling for Data Processing - Number Analytics, accessed on June 4, 2025, <https://www.numberanalytics.com/blog/10-stats-effective-min-max-scaling-data-processing>
10. arxiv.org, accessed on June 4, 2025, <https://arxiv.org/pdf/2502.09625>
11. Guaranteeing Out-Of-Distribution Detection in Deep RL via Transition Estimation - arXiv, accessed on June 4, 2025, <https://arxiv.org/html/2503.05238v1>
12. Time Series Anomaly Detection via Reinforcement Learning-Based Model Selection - arXiv, accessed on June 4, 2025, <https://arxiv.org/pdf/2205.09884>
13. the-signature-method-in-machine-learning/README.md at master - GitHub, accessed on June 4, 2025, <https://github.com/kormilitzin/the-signature-method-in-machine-learning/blob/master/README.md>
14. Higher Order Kernel Mean Embeddings to Capture Filtrations of Stochastic Processes - OpenReview, accessed on June 4, 2025, <https://openreview.net/pdf?id=CtugaUzfYw>
15. The Signature Method with Sktime, accessed on June 4, 2025, <https://www.sktime.net/en/latest/examples/transformation/signature_method.html>
16. A Primer on the Signature Method in Machine Learning - arXiv, accessed on June 4, 2025, <https://arxiv.org/pdf/1603.03788>
17. arxiv.org, accessed on June 4, 2025, <https://arxiv.org/html/2506.01718v1>
18. [2506.01718] Signature Maximum Mean Discrepancy Two-Sample Statistical Tests - arXiv, accessed on June 4, 2025, <https://arxiv.org/abs/2506.01718>
19. arxiv.org, accessed on June 4, 2025, <https://arxiv.org/html/2205.12706v4>
20. MMD-Sense-Analysis: Word Sense Detection Leveraging Maximum Mean Discrepancy, accessed on June 4, 2025, <https://arxiv.org/html/2506.01602v1>
21. maximum\_mean\_discrepancy - Kaggle, accessed on June 4, 2025, <https://www.kaggle.com/code/onurtunali/maximum-mean-discrepancy>
22. arxiv.org, accessed on June 4, 2025, <https://arxiv.org/html/2407.19848v3>
23. generative model for financial time series trained - arXiv, accessed on June 4, 2025, <https://arxiv.org/pdf/2407.19848>
24. Generative model for financial time series trained with MMD using a signature kernel - arXiv, accessed on June 4, 2025, <https://arxiv.org/abs/2407.19848>
25. issaz/signature-regime-detection: Code accompanying the paper "Pathwise methods for non-parametric online market regime detection and regime clustering for multidimensional and non-Markovian data" - GitHub, accessed on June 4, 2025, <https://github.com/issaz/signature-regime-detection>
26. maudl3116/higherOrderKME: Code for the NeurIPS 2021 paper "Higher Order Kernel Mean Embeddings to Capture Filtrations of Stochastic Processes". - GitHub, accessed on June 4, 2025, <https://github.com/maudl3116/higherOrderKME>
27. Universal Time-Series Representation Learning: A Survey - arXiv, accessed on June 4, 2025, <https://arxiv.org/html/2401.03717v1>
28. arxiv.org, accessed on June 4, 2025, <https://arxiv.org/pdf/2501.17992>
29. arXiv:2504.02281v3 [cs.CE] 16 May 2025, accessed on June 4, 2025, <https://www.arxiv.org/pdf/2504.02281>
30. Deep Convolutional Transformer Network for Stock Movement Prediction - MDPI, accessed on June 4, 2025, <https://www.mdpi.com/2079-9292/13/21/4225>
31. arxiv.org, accessed on June 4, 2025, <https://arxiv.org/html/2502.15813v1>
32. An attention embedded DUAL-LSTM method for financial risk early warning of the three new board-listed companies, accessed on June 4, 2025, <https://pmc.ncbi.nlm.nih.gov/articles/PMC10280675/>
33. LSTM–Transformer-Based Robust Hybrid Deep Learning Model for Financial Time Series Forecasting - MDPI, accessed on June 4, 2025, <https://www.mdpi.com/2413-4155/7/1/7>
34. Transformer vs. LSTM: 4 Key Differences and How to Choose - Kolena, accessed on June 4, 2025, <https://www.kolena.com/guides/transformer-vs-lstm-4-key-differences-and-how-to-choose/>
35. FinMTEB: Finance Massive Text Embedding Benchmark - arXiv, accessed on June 4, 2025, <https://arxiv.org/html/2502.10990v1>
36. arXiv:2502.10990v2 [cs.CL] 26 Feb 2025, accessed on June 4, 2025, <https://arxiv.org/pdf/2502.10990>
37. Beyond the Black Box: Interpretability of LLMs in FinanceThe views expressed in this paper are those of the authors and do not necessarily reflect the views of Barclays. - arXiv, accessed on June 4, 2025, <https://arxiv.org/html/2505.24650v1>
38. arxiv.org, accessed on June 4, 2025, <https://arxiv.org/html/2503.22192v1>
39. From Deep Learning to LLMs: A survey of AI in Quantitative Investment - arXiv, accessed on June 4, 2025, <https://arxiv.org/html/2503.21422v1>
40. Financial Time Series Forecasting with the Deep Learning Ensemble Model - MDPI, accessed on June 4, 2025, <https://www.mdpi.com/2227-7390/11/4/1054>
41. Forecasting stock prices with a feature fusion LSTM-CNN model using different representations of the same data, accessed on June 4, 2025, <https://pmc.ncbi.nlm.nih.gov/articles/PMC6377125/>
42. Backpropagation | Practical Reinforcement Learning for Robotics and AI, accessed on June 4, 2025, <https://www.reinforcementlearningpath.com/backpropagation/>
43. Real-Time Recurrent Reinforcement Learning - AAAI Publications, accessed on June 4, 2025, <https://ojs.aaai.org/index.php/AAAI/article/view/34001/36156>
44. Deep Reinforcement Learning for Quantitative Trading - arXiv, accessed on June 4, 2025, <https://arxiv.org/html/2312.15730v1>
45. ojs.aaai.org, accessed on June 4, 2025, <https://ojs.aaai.org/index.php/AAAI/article/view/16144/15951>
46. CaRL: Learning Scalable Planning Policies with Simple Rewards - arXiv, accessed on June 4, 2025, <https://arxiv.org/html/2504.17838v1>