# **פיתוח מערכת מסחר מרובת סוכנים מבוססת למידת חיזוק, נתוני רמה 2, זיהוי משטרים וסופרפוזיציה קוונטית למסחר בחוזים עתידיים**

## **פרק 1: יסודות מושגיים של מערכת המסחר מרובת הסוכנים**

### **1.1. סקירה כללית של המערכת המוצעת: שילוב MARL, נתוני רמה 2, זיהוי משטרים ומושגים בהשראת קוונטים למסחר בחוזים עתידיים**

היוזמה המונחת ביסוד מחקר זה עוסקת בתכנון והקמה של מערכת מסחר מתוחכמת ביותר, המיועדת לפעול בשוק החוזים העתידיים. מערכת זו שואפת למזג באופן סינרגטי מספר טכנולוגיות מתקדמות מתחום הבינה המלאכותית ומדעי הנתונים, במטרה להשיג יתרון תחרותי בסביבת המסחר הדינמית. הרעיון המרכזי הוא פיתוח מערכת למידת חיזוק מרובת סוכנים (MARL), אשר תוכננה לפעול בשיתוף פעולה הדוק. ייחודה של המערכת טמון בשילוב של מספר מרכיבים מרכזיים: ניצול נתוני שוק מסוג רמה 2 (Level 2 data) בזמן אמת, מנוע זיהוי משטרי שוק המבוסס על מדד Higher-Rank Maximum Mean Discrepancy (MMD), ולבסוף, שילוב עקרונות בהשראת מכניקת הקוונטים, ובפרט רעיון הסופרפוזיציה, לשם שיפור עמידות המערכת.

כל אחד ממרכיבים אלו תורם היבט ייחודי לתפקודה הכולל של המערכת. נתוני רמה 2 מספקים תמונה עשירה ומפורטת של עומק השוק וספר הפקודות, ומאפשרים לסוכנים לקבל החלטות מסחר מבוססות מידע מיקרו-מבני. מערכת ה-MARL מאפשרת עיבוד מבוזר של נתונים אלו על ידי הקצאת מקטעי נתונים שונים לסוכנים בודדים, תוך שמירה על כוונת אסטרטגית מאוחדת באמצעות מדיניות משותפת. מנוע זיהוי משטרי השוק באמצעות MMD מסדר גבוה נועד לספק לסוכנים מודעות הקשרית רחבה יותר לגבי המצב המאקרו-כלכלי של השוק, ולאפשר התאמה של האסטרטגיה למשטר השוק הנוכחי. לבסוף, הרעיון של סופרפוזיציה בהשראת קוונטים נחקר כאמצעי להגברת העמידות והאמינות של הסוכנים הבודדים ושל המערכת כולה, על ידי מתן אפשרות לסוכנים לשקול מספר מצבים או פעולות בו-זמנית באופן הסתברותי.

הארכיטקטורה המוצעת שואפת לאזן בין עיבוד נתונים מבוזר, המתבצע על ידי סוכנים בודדים הפועלים על מקטעי נתונים ייחודיים להם, לבין אינטליגנציה אסטרטגית מרכזית, המגולמת במדיניות המסחר המשותפת ובמודעות למשטר השוק. הצלחתה של גישה זו תלויה ביכולת ליצור הקשר משמעותי מספיק עבור כל סוכן בודד מתוך מקטע הנתונים שלו, ובאפקטיביות של מנגנון שיתוף הפעולה להפקת התנהגות אינטליגנטית קולקטיבית. מערכת כזו, אם תוכח כיעילה, עשויה להוות תבנית חדשנית לבניית מערכות מסחר מתוחכמות המסוגלות להסתגל בו-זמנית להיבטים שונים של השוק, ולהוביל לקבלת החלטות מסחר חזקות ומדויקות יותר מאשר אלו שניתן להשיג באמצעות סוכן יחיד ומונוליטי.

### **1.2. הבנת תשומות מפתח: חוזים עתידיים ונתוני רמה 2 של רית'מיק**

הבנה מעמיקה של תשומות הליבה – חוזים עתידיים ונתוני רמה 2 – חיונית, שכן אלו מהווים את חומר הגלם ואת כלי המסחר עבור מערכת ה-MARL המוצעת.

**חוזים עתידיים (Futures Contracts):** חוזה עתידי הוא מכשיר פיננסי המאפשר גידור סיכונים או ספקולציה על מחירי נכסים עתידיים. חוזים אלו הם סטנדרטיים ונסחרים בבורסות ייעודיות, והם מוגדרים על ידי תנאים קבועים מראש כגון סוג המוצר, מחירו, ומועד ביצוע החוזה. הנכסים הבסיסיים לחוזים עתידיים יכולים להיות מגוונים, וכוללים סחורות, מטבעות חוץ, ומדדי מניות. מערכת המסחר המוצעת תתמקד ככל הנראה בחוזים עתידיים הנסחרים בבורסה (Futures) ולא בחוזי Forward, שהם מותאמים אישית ונסחרים מעבר לדלפק (OTC). התכונות של סטנדרטיזציה, נזילות גבוהה יחסית, והיכולת למנף הופכות את החוזים העתידיים למכשיר מתאים עבור מערכות מסחר אלגוריתמיות מתקדמות.

**נתוני רמה 2 (Level 2 Data):** נתוני רמה 2 מספקים תמונה מפורטת של ספר הפקודות בזמן אמת, מעבר למחיר ההצעה (bid) והדרישה (ask) הטובים ביותר (NBBO). נתונים אלו כוללים את כל הפקודות הפתוחות (limit orders) במחירים שונים מתחת להצעה הטובה ביותר ומעל לדרישה הטובה ביותר, יחד עם כמויות הפקודות בכל רמת מחיר. מידע זה מכונה גם "עומק שוק" (market depth) או "ספר פקודות" (order book). ספקי נתונים כמו רית'מיק (Rithmic), שהוזכרו בהקשר של המערכת, יכולים להציג את מלוא עומק השוק כפי שהוא מוצג על ידי הבורסה.

ניתוח נתוני רמה 2 מאפשר לסוחרים להעריך את היצע וביקוש לנכס מסוים, לזהות רמות תמיכה והתנגדות פוטנציאליות על סמך ריכוזי פקודות גדולים, ולהעריך את נזילות השוק ברמות מחיר שונות. מידע זה קריטי במיוחד למסחר בתדירות גבוהה (HFT) או לאסטרטגיות הרגישות למיקרו-מבנה השוק. עם זאת, חשוב לציין כי לא כל הפקודות מופיעות בנתוני רמה 2; פקודות גדולות של משקיעים מוסדיים עשויות להיות מוסתרות או מחולקות (פקודות קרחון - iceberg orders). בנוסף, תנועות מחירים הנצפות בנתוני רמה 2 מייצגות נזילות זמינה ולאו דווקא עסקאות שבוצעו בפועל, ויכולות להיות מושפעות מפעילות של תוכנות HFT. העושר והמורכבות של נתוני רמה 2 הם אלו שמערכת ה-MARL תנסה לנצל.

### **1.3. פרדיגמת למידת חיזוק מרובת סוכנים (MARL) בפיננסים**

למידת חיזוק מרובת סוכנים (MARL) מהווה הרחבה של למידת חיזוק (RL) של סוכן יחיד, ומתייחסת למצבים בהם מספר סוכנים לומדים ומקבלים החלטות באופן סימולטני בסביבה משותפת. בניגוד ל-RL של סוכן יחיד, שם הסוכן לומד מדיניות אופטימלית באמצעות אינטראקציה עם סביבה סטטית יחסית, ב-MARL כל סוכן צריך לקחת בחשבון את פעולותיהם והשפעתם של סוכנים אחרים, מה שהופך את הסביבה לדינמית ולא-סטציונרית מנקודת מבטו של כל סוכן בודד. מערכות MARL מסווגות לעיתים קרובות על בסיס פונקציות התגמול שלהן לשלוש קטגוריות: שיתופיות מלאה, תחרותיות מלאה, או מעורבות.

הפוטנציאל של MARL בתחום הפיננסי נובע מיכולתה למדל אינטראקציות מורכבות בשוק ולפתח אסטרטגיות שיתופיות או תחרותיות. מחקרים בתחום ה-MARL בוחנים מדדים חברתיים כגון שיתוף פעולה, הדדיות והשפעה חברתית, שהם רלוונטיים לדרישת "שיתוף הפעולה המושלם" במערכת המוצעת. יישומים של MARL בפיננסים כוללים אופטימיזציה של אסטרטגיות מסחר בתדירות גבוהה וניהול תיקי השקעות , כאשר מחקרים מראים כי גישות MARL יכולות להניב ביצועים עדיפים על פני מודלים של סוכן יחיד.

עם זאת, MARL מציבה אתגרים משמעותיים, ביניהם בעיית הממדיות (curse of dimensionality) הנובעת ממרחבי מצב-פעולה משותפים גדולים, אי-סטציונריות הסביבה, תצפית חלקית (partial observability) של כל סוכן, וסקלביליות המערכת למספר גדול של סוכנים. המערכת המוצעת, עם הדגש שלה על שיתוף פעולה מושלם וחלוקת נתונים בין סוכנים, תצטרך להתמודד ישירות עם אתגרים אלו. הבחירה ב-MARL מונעת מהרצון לבזר את עיבוד נתוני רמה 2 המורכבים תוך שמירה על גישה אסטרטגית מאוחדת, כאשר האתגר המרכזי יהיה להבטיח שהאינטראקציה בין הסוכנים תוביל להתנהגות אינטליגנטית קולקטיבית ולא רק לפעולות מקבילות אך לא מתואמות.

## **פרק 2: צנרת הנתונים: מרכישת נתונים גולמיים ועד למטריצות קלט ספציפיות לסוכן**

צנרת הנתונים מהווה נדבך קריטי במערכת המסחר המוצעת. היא אחראית על הפיכת נתוני שוק גולמיים לפורמט מתאים ועתיר מידע, אשר ישמש כקלט עבור סוכני ה-MARL. תהליך זה כולל רכישה ועיבוד של נתוני רמה 2, זיהוי משטרי שוק באמצעות טכניקות MMD מתקדמות, ושילוב מידע זה טרם חלוקת הנתונים למטריצות ייעודיות עבור כל סוכן.

### **2.1. רכישה ופירוש של נתוני רמה 2 מרית'מיק**

כפי שצוין קודם לכן (פרק 1.2), נתוני רמה 2 מספקים תמונה מפורטת של ספר הפקודות, כולל הצעות מחיר (bids) ודרישות (asks) מעבר למחירים הטובים ביותר, יחד עם כמויות הפקודות בכל רמת מחיר. עבור המערכת המוצעת, נתונים אלה יירכשו דרך ספק הנתונים רית'מיק, הידוע ביכולתו לספק את מלוא עומק השוק (Depth of Market).

איכות הנתונים היא בעלת חשיבות עליונה. יש להבטיח חותמות זמן מדויקות, טיפול עקבי בנתונים חסרים (אם קיימים), והבנה של מגבלות פוטנציאליות. למשל, פקודות מוסתרות (hidden orders) או פקודות קרחון (iceberg orders), בהן פקודה גדולה מחולקת לבלוקים קטנים יותר המוזרמים לשוק בהדרגה, עשויות שלא להשתקף במלואן בנתוני רמה 2. תופעה זו, יחד עם האפשרות שתוכנות מסחר בתדירות גבוהה (HFT) משפיעות על נתוני רמה 2 גם ללא ביצוע עסקאות בפועל , מדגישה את הצורך בגישה ביקורתית וזהירה בעת פירוש הנתונים.

נתוני רמה 2 מגיעים כזרם של עדכונים, וניתן לארגנם במגוון דרכים לצורך ניתוח והנדסת מאפיינים. למשל, ניתן לקבץ עדכונים לפי סוג הפקודה (קנייה/מכירה), לפי סוג הפעולה (הוספה/ביטול פקודה), לפי המרחק מהמחיר הטוב ביותר, או לפי גודל הפקודה. גמישות זו בארגון הנתונים היא קריטית לשלב הנדסת המאפיינים, בו יחולצו תכונות רלוונטיות עבור סוכני ה-RL.

### **2.2. מנוע זיהוי משטרי שוק באמצעות MMD מסדר גבוה**

זיהוי משטרי שוק (market regime detection) הוא תהליך קריטי עבור אסטרטגיות מסחר רבות, שכן ביצועי הנכסים והאסטרטגיות עשויים להשתנות באופן משמעותי בין משטרים שונים (למשל, שוק שורי, שוק דבי, שוק תנודתי, שוק מגמתי). המערכת המוצעת תכלול מנוע ייעודי לזיהוי משטרי שוק, המבוסס על מדד Higher-Rank Maximum Mean Discrepancy (MMD).

#### **2.2.1. עקרונות MMD ו-MMD מסדר גבוה לזיהוי משטרי שוק**

Maximum Mean Discrepancy (MMD) הוא מבחן סטטיסטי לא-פרמטרי המשמש להערכת הדמיון בין שתי התפלגויות הסתברות, על סמך דגימות מאותן התפלגויות. הרעיון המרכזי הוא להטמיע את ההתפלגויות במרחב מאפיינים בעל ממד גבוה (reproducing kernel Hilbert space - RKHS) באמצעות פונקציית גרעין (kernel function), ולמדוד את המרחק בין הייצוגים הממוצעים של ההתפלגויות במרחב זה. אם המרחק גדול מספיק, ניתן להסיק שההתפלגויות שונות.

בהקשר של זיהוי משטרי שוק, MMD יכול לשמש להשוואת התפלגות נתוני השוק (למשל, תשואות, תנודתיות) בחלון זמן נוכחי להתפלגות בחלון זמן היסטורי המשמש כרפרנס. שינוי משמעותי בהתפלגות יכול להצביע על מעבר למשטר שוק חדש. המחקר של Horvath ו-Issa (2023) מציג גישה מתקדמת המשתמשת במבחן דגימה דו-מסלולי (path-wise two-sample test) המבוסס על מדד דמיון מסוג MMD על מרחב מסלולים (path space). גישה זו משתמשת בחתימות נתיבים מחוספסים (rough path signatures) כמפת מאפיינים (feature map). חתימות אלו מסוגלות ללכוד מידע על סדר התצפיות ועל התלות ביניהן לאורך זמן, מה שהופך אותן למתאימות במיוחד לניתוח סדרות עתיות פיננסיות, שהן לרוב לא-מרקוביות.

המונח "MMD מסדר גבוה" (Higher-Rank MMD), כפי שמוזכר ב-, מתייחס ליכולת לשלב מידע פילטרציה (filtration information) בתהליך ההסקה. במילים אחרות, המדד לוקח בחשבון את התפתחות המידע לאורך זמן, ולא רק התפלגויות סטטיות. זהו יתרון משמעותי בניתוח שווקים פיננסיים, שבהם ההיסטוריה והסדר של האירועים חשובים. הגישה המוצעת ב- שואפת להיות ישימה בזמן אמת (online) ולספק אינדיקציה מוקדמת לשינויים התפלגותיים, גם אם הם קטנים או זמניים. יתרונות נוספים של גישת sig-MMD כוללים הימנעות מצורך בקיטוע (truncation) של החתימות ויכולת התמודדות עם נתונים רב-ממדיים.

טכניקות נוספות לזיהוי משטרי שוק, כגון מבחן קולמוגורוב-סמירנוב, מרחק וסרשטיין, CUSUM ו-DTW, יכולות לספק הקשר נוסף ורעיונות לגבי פרמטרים רלוונטיים כמו אורכי חלונות וספים. חשוב לציין את הרעיון של שילוב מספר אינדיקטורים ובדיקתם על פני מסגרות זמן מרובות לשם אישוש האותות.

#### **2.2.2. אסטרטגיית יישום: חילוץ פרמטרי משטר ושילובם בזרם הנתונים**

הפלט של מנוע זיהוי המשטירים באמצעות MMD מסדר גבוה יכול להיות מדד מרחק (המצביע על מידת השוני בין התקופה הנוכחית לתקופת הרפרנס) או תווית קטגוריאלית המציינת את משטר השוק הנוכחי (למשל, "תנודתיות גבוהה", "מגמת עלייה חזקה"). פרמטרים אלו, או מאפיינים נוספים הנגזרים מהם (כגון רמת תנודתיות צפויה במשטר הנוכחי, עוצמת המגמה), ישולבו כמאפיינים נוספים בכל נקודת נתונים (למשל, כל תמונת מצב של ספר הפקודות מרמה 2).

השילוב של פרמטרי המשטר בנתונים יתבצע *לפני* שלב החלוקה האקראית למטריצות. כך, כל סוכן, למרות שהוא מקבל רק מקטע חלקי של נתוני השוק, יהיה מודע להקשר המאקרו-כלכלי הרחב יותר שמספק מנוע זיהוי המשטירים. מאגר הקוד issaz/signature-regime-detection ב-GitHub מכיל קוד פייתון הרלוונטי ליישום שיטות אלו, ובפרט המחברות 4-online-regime-detection.ipynb ו-4-higher-rank-mmd.ipynb הן מקורות מרכזיים לפרטי המימוש. חבילת higherOrderKME מוזכרת כחיונית לחישובי signature kernel MMD. בעוד שדוגמאות קוד אחרות משתמשות במודלים כמו HMM או GMM לזיהוי משטרים, זרימת העבודה של אימון מודל לזיהוי משטרים ושימוש בפלט שלו כמאפיין נוסף היא אנלוגית.

הטבלה הבאה מסכמת היבטים מרכזיים של שימוש ב-MMD מסדר גבוה לזיהוי משטרי שוק:

**טבלה 1: סקירת טכניקות MMD מסדר גבוה לזיהוי משטרי שוק**

| **מאפיין/מושג MMD** | **עיקרון (רעיון מתמטי/סטטיסטי)** | **פרמטרים מרכזיים** | **יתרונות לנתונים פיננסיים** | **אתגרים/מגבלות** | **הערות יישום/ספריות פייתון** | **מקורות רלוונטיים** |
| --- | --- | --- | --- | --- | --- | --- |
| MMD כללי | מבחן לא-פרמטרי להשוואת התפלגויות באמצעות הטבעה ל-RKHS ומדידת מרחק בין ממוצעים. | פונקציית גרעין (למשל, גאוסיאני, לינארי), פרמטרי הגרעין (למשל, רוחב פס). | גמישות, אי-תלות בהנחות התפלגותיות. | בחירת גרעין ופרמטרים אופטימליים, רגישות לממדיות גבוהה. | scikit-learn (לגרעינים), ספריות ייעודיות ל-MMD. |  |
| Signature MMD (sig-MMD) | שימוש בחתימות נתיבים מחוספסים (rough path signatures) כמפת מאפיינים עבור MMD, לוכד מידע סדרתי ותלותי בזמן. | רמת קיטוע החתימה , סוג הגרעין על מרחב החתימות. | טיפול טבעי בסדרות עתיות, לכידת תלויות לא-מרקוביות, אינווריאנטיות לפרמטריזציה מחדש של הזמן. | מורכבות חישובית של חתימות מסדר גבוה, פרשנות של מרחב החתימות. | iisignature (לחישוב חתימות), higherOrderKME (ל-signature kernel MMD). |  |
| Higher-Rank MMD | הרחבה של MMD המאפשרת שילוב מידע פילטרציה (התפתחות המידע לאורך זמן) בתהליך ההסקה. | הגדרת "דרגות" או רמות פילטרציה, בחירת גרעינים מתאימים לכל דרגה. | מתאים במיוחד לסביבות לא-מרקוביות כמו שווקים פיננסיים, מאפשר ניתוח עשיר יותר של דינמיקות התפלגותיות. | מורכבות מושגית וחישובית גבוהה יותר, דרישות נתונים פוטנציאליות גבוהות יותר לאמידת התפלגויות מסדר גבוה. | קוד ייעודי כפי שמוצע במחקר של Horvath & Issa , עשוי לדרוש פיתוח מותאם אישית. |  |
| Online Regime Detection | יישום MMD (לרוב sig-MMD) באופן מקוון להשוואת חלון נתונים עדכני לחלון רפרנס היסטורי, במטרה לזהות שינויים התפלגותיים בזמן אמת. | גודל חלון ההיסטוריה, גודל חלון הנתונים החדשים, סף סטטיסטי להכרזה על שינוי משטר, תדירות עדכון. | יכולת תגובה מהירה לשינויים בשוק, זיהוי מוקדם של מעברי משטר. | איזון בין רגישות (זיהוי שינויים קטנים) לבין יציבות (הימנעות מאזעקות שווא), עיכוב בזיהוי (תלוי בגודל החלון). | מאגר הקוד issaz/signature-regime-detection , ובפרט המחברת 4-online-regime-detection.ipynb. |  |

### **2.3. בניית מבנה נתונים עבור MARL**

לאחר רכישת נתוני רמה 2 והעשרתם בפרמטרי משטר השוק, השלב הבא הוא ארגון הנתונים באופן שיתאים לקלט של מערכת ה-MARL. שלב זה כולל את החלוקה האקראית של הנתונים למטריצות והנדסת מאפיינים ספציפית לכל מטריצה.

#### **2.3.1. חלוקה אקראית של נתונים מודעי-משטר למטריצות (Ma​,Mb​,...,Mn​)**

הרעיון המרכזי שהוצג הוא "להפריד באופן אקראי את הנתונים ולארגן אותם למטריצות (Ma​,Mb​,...,Mn​), כאשר כל מטריצה מייצגת מקטעים קטנים יותר של הנתונים, וכל סוכן RL יופנה לכל מטריצה". גישה זו נועדה ככל הנראה להשיג מספר מטרות: (א) הקטנת ממדיות הקלט עבור כל סוכן בודד, (ב) מתן אפשרות לסוכנים "להתמחות" בהיבטים שונים או "תצוגות" (views) שונות של נתוני השוק, ו-(ג) לאפשר עיבוד מקבילי.

אופן החלוקה ה"אקראית" דורש הבהרה. האם מדובר בדגימה אקראית של רשומות מספר הפקודות? האם מדובר בחלוקה אקראית של מקטעים זמניים? או שמא בחלוקה אקראית של תת-קבוצות של מאפיינים הנגזרים מספר הפקודות? ההנחיה מציינת "מקטעים קטנים יותר של הנתונים", מה שמרמז על חלוקה של זרם הנתונים או של מאפייניו. מספר המטריצות (N) גמיש, וניתן להתייחס אליו כהיפר-פרמטר של המערכת.

חשוב לוודא שכל מטריצה (Ma​,Mb​,...) עדיין מכילה מידע משמעותי מספיק עבור סוכן RL. חלוקה מוגזמת או יצירת מקטעים דלי-הקשר עלולה לפגוע ביכולת הלמידה של הסוכנים. בעוד שמחקרים קיימים דנים בייצוג מצב כללי ב-RL ובאתגרי MARL כמו ממדיות ותצפית חלקית , הם אינם מפרטים אסטרטגיות חלוקת נתונים פיננסיים ספציפיות כפי שמוצע כאן. עם זאת, הרעיון של סוכנים בעלי תצוגות נתונים שונות או מתמחים נתמך על ידי מחקרים המשלבים מקורות נתונים מרובים או טכניקות הפחתת ממדים באמצעות הטלות אקראיות. מחקרים אחרים ב-MARL עוסקים בהקצאת תכונות שונות לסוכנים שונים או בקידוד מידע ממספר סוכנים לייצוג מצב מאוחד.

האתגר המרכזי בחלוקה אקראית זו הוא האיזון בין יצירת "תצוגות" מגוונות עבור הסוכנים השונים לבין שמירה על איכות ועומק המידע בכל תצוגה. פרמטרי המשטר שנוספו בשלב הקודם עשויים לסייע במתן הקשר גלובלי גם לתצוגות חלקיות אלו. עם זאת, הצלחת המערכת תהיה תלויה במידה רבה ביכולת להגדיר אסטרטגיית חלוקה שתשמר מספיק מידע מקומי בכל מטריצה, כך שהסוכנים יוכלו לפעול באופן קוהרנטי במסגרת האסטרטגיה הגלובלית המשותפת.

#### **2.3.2. הנדסת מאפיינים עבור מטריצות הקלט של הסוכנים מתוך מקטעי נתוני רמה 2**

לאחר שנוצר "מקטע" נתונים (Mi​) עבור סוכן ספציפי, יש להפיק ממנו מאפיינים (features) שייצגו את מצב הסביבה עבור אותו סוכן. אופי המאפיינים יהיה תלוי באופן בו הוגדר המקטע:

* **אם המקטע הוא תת-קבוצה של רמות מחיר בספר הפקודות:** המאפיינים יכולים לכלול מחירי היצע וביקוש, כמויות, הפרשים ביניהם, וחוסר איזון מקומי ברמות אלו.
* **אם המקטע הוא קטע זמני:** המאפיינים יכולים לכלול רצפים של אירועים בספר הפקודות, סטטיסטיקות מחיר/כמות על פני אותו פרק זמן קצר, וקצב הגעת פקודות.
* **אם המקטע הוא תת-קבוצה אקראית של כלל מאפייני רמה 2 הזמינים:** זוהי הגישה הישירה ביותר, אך דורשת בחירה מושכלת של המאפיינים הגולמיים.

מקורות כמו מציינים שניתן לקבץ עדכוני רמה 2 לפי קנייה/מכירה, הוספה/ביטול פקודה, מרחק מהמחיר הטוב ביותר, וגודל עסקה – כל אלו יכולים להוות בסיס ליצירת מאפיינים. מחקרים על RL למסחר אופטימלי משתמשים במאפיינים כמו כמות ונפח ברמות מחיר שונות בספר הפקודות. פרמטרי המשטר שזוהו בשלב 2.2.2 חייבים להיכלל במערך המאפיינים של כל סוכן, כדי לספק את ההקשר המאקרו-כלכלי.

הטבלה הבאה מציגה דוגמאות למאפיינים פוטנציאליים שניתן להנדס מנתוני רמה 2 ומפרמטרי המשטר עבור קלט הסוכנים:

**טבלה 2: מאפייני נתוני רמה 2 עבור מטריצות קלט של סוכנים**

| **קטגוריית מאפיין** | **דוגמאות למאפיינים ספציפיים** | **שיטת גזירה/נוסחה** | **תועלת פוטנציאלית לקבלת החלטות של סוכן RL** | **מקורות רלוונטיים** |
| --- | --- | --- | --- | --- |
| מצב ספר פקודות גולמי (LOB) | מחירי היצע/ביקוש ברמה X, כמויות היצע/ביקוש ברמה X, מרווח היצע-ביקוש (spread) ברמה X, עומק מצטבר עד רמה X. | קריאה ישירה מהמקטע Mi​. | מספק תמונה של מבנה השוק המיידי הרלוונטי למקטע הנתונים של הסוכן. |  |
| זרימת פקודות (Order Flow) | קצב הגעת פקודות חדשות (limit/market), קצב ביטול פקודות, יחס בין פקודות קנייה למכירה. | חישוב סטטיסטי על פני חלון זמן קצר בתוך המקטע Mi​. | מזהה לחצי קנייה/מכירה מיידיים ושינויים במומנטום. |  |
| חוסר איזון (Imbalance) | חוסר איזון בין היצע לביקוש ברמות מחיר קרובות, חוסר איזון משוקלל נפח (VWAP imbalance). | חישוב יחסים או הפרשים בין כמויות היצע וביקוש. | יכול להצביע על כיוון תנועת מחיר צפוי בטווח הקצר. |  |
| תנודתיות (Volatility) | תנודתיות היסטורית מחושבת על המקטע Mi​, תנודתיות נרמזת (אם זמינה ורלוונטית למקטע). | סטיית תקן של שינויי מחירים, או מדדים מורכבים יותר. | הערכת רמת הסיכון הנוכחית והתאמת גודל הפוזיציה או אגרסיביות המסחר. |  |
| נזילות (Liquidity) | סך הכמות הזמינה ב-X רמות המחיר הקרובות, זמן ממוצע לביצוע פקודה (אם ניתן להעריך מהנתונים). | סיכום כמויות מספר הפקודות. | הערכת היכולת להיכנס ולצאת מפוזיציות בגודל הרצוי ללא השפעה משמעותית על המחיר. |  |
| אינדיקטורים ממשטר MMD | תווית משטר נוכחי (למשל, 1=מגמתי, 2=תנודתי), פרמטר תנודתיות משויך למשטר, עוצמת מגמה משוערת. | פלט ישיר ממנוע זיהוי המשטירים (פרק 2.2). | מספק הקשר מאקרו להתנהגות השוק, ומאפשר לסוכן להתאים את יישום האסטרטגיה המשותפת למצב השוק הכולל. |  |
| מאפיינים מבוססי זמן/אירועים | זמן מאז העדכון האחרון בספר הפקודות, מספר עדכונים ב-X השניות האחרונות, גודל עסקה ממוצע אחרון. | מעקב אחר חותמות זמן וספירת אירועים. | לכידת דינמיקות מהירות ומיקרו-מבנה של השוק. |  |

המורכבות של צנרת הנתונים, החל מרכישת נתוני רמה 2, דרך זיהוי משטרים מתוחכם באמצעות MMD, ועד לחלוקה אקראית והנדסת מאפיינים עבור סוכני MARL, מדגישה את האתגר הטמון במימוש המערכת. הצלחת המערכת תלויה באופן קריטי ביכולת של כל שלב בצנרת לספק מידע איכותי, רלוונטי ובעל הקשר מספק לשלב הבא. החלוקה האקראית של הנתונים, בפרט, מהווה נקודה קריטית הדורשת תכנון זהיר כדי להבטיח שהסוכנים הבודדים אכן מקבלים תמונה משמעותית מספיק של השוק, גם אם חלקית, כדי שיוכלו ליישם את האסטרטגיה המשותפת באופן אפקטיבי. פרמטרי המשטר הגלובליים אמורים לסייע בכך, אך האיזון בין פירוט מקומי להקשר גלובלי יוותר אתגר מרכזי.

## **פרק 3: תכנון מערכת MARL שיתופית**

תכנון מערכת ה-MARL השיתופית עומד בלב הפרויקט, ודורש התייחסות למספר היבטים מרכזיים: ארכיטקטורת הליבה של הסוכנים, אופן ייצוג וביצוע המדיניות המשותפת, הבטחת שיתוף פעולה אפקטיבי, והגדרת מרחבי המצב והפעולה של הסוכנים הבודדים. כל זאת, תוך מינוף ספריות וכלים קיימים בתחום ה-MARL.

### **3.1. ארכיטקטורת MARL מרכזית למסחר מתואם**

בהינתן הדרישה ל"שיתוף פעולה מושלם" ולמדיניות מסחר משותפת, פרדיגמת אימון מרכזי עם ביצוע מבוזר (Centralized Training with Decentralized Execution - CTDE) נראית כמתאימה ביותר. בפרדיגמה זו, במהלך שלב האימון, ניתן להשתמש במידע גלובלי (למשל, מצבים ופעולות של כל הסוכנים, או פונקציית תגמול גלובלית) כדי לאמן את מדיניות הסוכנים או פונקציית ערך מרכזית. במהלך שלב הביצוע (מסחר חי או בדיקות לאחור), כל סוכן פועל באופן מבוזר על סמך התצפית המקומית שלו (מטריצת הנתונים Mi​ ופרמטרי המשטר) והמדיניות המשותפת שנלמדה.

מערכות MARL נועדו להתמודד עם אינטראקציות ותלות בין סוכנים במטרה להשיג אופטימיזציה גלובלית. במקרה שלנו, האופטימיזציה הגלובלית מוגדרת על ידי הצלחת אסטרטגיית המסחר הכוללת, המורכבת מארבע הסינרגיות.

#### **3.1.1. התמחות סוכנים מבוססת על חלוקות נתונים לעומת ביצוע מדיניות משותפת**

חשוב להדגיש כי במערכת המוצעת, ה"התמחות" של הסוכנים אינה נובעת ממדיניות שונה או מתפקידים שונים המוקצים להם, אלא מהעובדה שכל סוכן מקבל תצפית ייחודית (מקטע נתונים Mi​) על הסביבה. כל הסוכנים חולקים ומבצעים את אותה מדיניות מסחר. לפיכך, ההתמחות היא ברמת התפיסה (perception) ולא ברמת קבלת ההחלטות הלוגית. כל סוכן מיישם את אותה לוגיקה אסטרטגית, אך על בסיס מידע חלקי ושונה. זה דומה למצב בו מספר אנליסטים באותו צוות משתמשים באותה מתודולוגיית ניתוח, אך כל אחד מהם בוחן סט נתונים אחר כדי להגיע לתובנה קולקטיבית.

#### **3.1.2. הבטחת "שיתוף פעולה מושלם" בין סוכנים**

המונח "שיתוף פעולה מושלם" מחייב שהפעולות של כל הסוכנים יהיו קוהרנטיות ותורמות למטרת הצוות המשותפת, כפי שהיא מוגדרת על ידי האסטרטגיה והסינרגיות שלה. ניתן להשיג זאת במספר דרכים:

1. **מדיניות משותפת ופונקציית תגמול גלובלית:** הדרך הפשוטה ביותר היא להניח שהמדיניות המשותפת, כאשר היא מאומנת עם פונקציית תגמול גלובלית (למשל, רווח והפסד כולל של המערכת), תלמד באופן אימפליציטי את התיאום הנדרש. כל סוכן, הפועל למקסם את התגמול הגלובלי על סמך התצפית המקומית שלו, יתרום למאמץ המשותף.
2. **מנגנוני תקשורת (Communication Protocols):** אם תיאום אימפליציטי אינו מספיק, ניתן לשקול הוספת ערוצי תקשורת בין הסוכנים. סוכנים יכולים לשתף מידע רלוונטי מהתצפיות המקומיות שלהם, כוונות פעולה, או הערכות מצב. לדוגמה, מחקר מציע פרוטוקול תקשורת למידה רב-סיבובי, בו סוכנים מתקשרים לגבי כוונות הפעולה שלהם ומעדכנים אותן בהתאם, כדי לשפר שיתוף פעולה בביצוע פקודות מרובות. ספריות כמו CommNet ו-IC3Net מספקות מנגנונים כאלה.
3. **אסטרטגיות החלטה קולקטיביות:** ניתן להשתמש בפרוטוקולי הצבעה, אופטימיזציה מבוססת תועלת, או פונקציות תגמול משותפות כדי לגבש החלטה קולקטיבית.

האתגר המרכזי בשיתוף פעולה, כאשר סוכנים פועלים על סמך תצפיות חלקיות ואקראיות, הוא כיצד להבטיח שפעולותיהם, הנגזרות מאותה מדיניות, יהיו אכן מתואמות ולא סותרות ברמה הגלובלית. ייתכן שהמדיניות המשותפת תצטרך להיות מותנית גם על רכיב מצב גלובלי כלשהו (מעבר לפרמטרי המשטר), או שמבנה התגמול יצטרך להעניש בחריפות חוסר תיאום.

### **3.2. אסטרטגיית המסחר כמדיניות מאוחדת**

אסטרטגיית המסחר, הכוללת ארבע סינרגיות שונות עבור פוזיציות לונג ושורט, תהווה את ליבת המדיניות (π) של כל סוכני ה-RL.

#### **3.2.1. ייצוג ארבע הסינרגיות (לונג/שורט) במדיניות ה-RL**

יש להגדיר במדויק את "ארבע הסינרגיות" ואת התנאים להפעלתן. רשת המדיניות של סוכני ה-RL צריכה להיות מסוגלת לייצג ולבצע את הסינרגיות הללו. קיימות מספר אפשרויות לייצוג כזה:

* **פלט מדיניות הוליסטי:** רשת המדיניות יכולה להפיק פעולה אחת (למשל, כמות לקנייה/מכירה, מחיר פקודה) המשקללת או משלבת באופן אינהרנטי את כל ארבע הסינרגיות בהתאם למצב הקלט.
* **בחירת תת-מדיניות:** רשת המדיניות יכולה לבחור אחת מבין ארבע (או שמונה, אם לונג/שורט נפרדים לכל סינרגיה) תת-מדיניויות ספציפיות לסינרגיה, ולאחר מכן לבצע את הפעולה המומלצת על ידי תת-המדיניות שנבחרה.
* **התניית מצב:** מרחב המצב של הסוכנים יכול לכלול מאפיינים המצביעים על התאמה של סינרגיה מסוימת, והמדיניות תלמד להגיב באופן שונה למאפיינים אלו, ובכך להפעיל את הסינרגיה המתאימה.

המימוש ידרוש הבנה עמוקה של מהות הסינרגיות וכיצד הן מתורגמות ללוגיקה חישובית או לתבניות נתונים הניתנות לזיהוי על ידי רשת המדיניות.

#### **3.2.2. עיצוב מרחב הפעולות עבור הסוכנים**

מרחב הפעולות של הסוכנים צריך לשקף את הפעולות האפשריות במסחר בחוזים עתידיים, בהתאם לאסטרטגיה ולאופן ייצוג הסינרגיות. פעולות טיפוסיות יכולות לכלול:

* **סוג הפקודה:** פקודת שוק (market order), פקודת גבול (limit order).
* **כיוון הפוזיציה:** כניסה לפוזיציית לונג, כניסה לפוזיציית שורט, סגירת פוזיציה קיימת, או החזקה (ללא שינוי).
* **גודל הפוזיציה/הפקודה:** מספר החוזים לקנייה/מכירה.
* **מחיר הפקודה (עבור פקודות גבול):** רמת המחיר בה תוצב הפקודה, ייתכן ביחס לרמות מחיר קיימות בספר הפקודות (למשל, הצעה/דרישה הטובה ביותר + X טיקים).
* **ביטול/שינוי פקודות קיימות.**

דוגמאות ממחקרים כוללות מרחבי פעולה פשוטים כמו לונג/שורט/החזקה , או מרחבים מורכבים יותר המגדירים את מספר המניות לסחור , ואף בחירת סוג הפקודה והצבתה ברמות שונות בספר הפקודות. מרחב הפעולה צריך להיות מוגדר כך שיאפשר ביטוי מלא של ארבע הסינרגיות.

### **3.3. ייצוג מצב עבור סוכנים בודדים (שילוב נתונים מחולקים והקשר משטרי)**

כל סוכן i במערכת ה-MARL מקבל כקלט את מטריצת הנתונים הייחודית לו, Mi​ (הנגזרת מהחלוקה האקראית של נתוני רמה 2), יחד עם פרמטרי משטר השוק הגלובליים שזוהו על ידי מנוע ה-MMD. המצב si​ עבור סוכן i יהיה פונקציה של שני מקורות מידע אלו. יש להמיר קלט זה לייצוג וקטורי/טנזורי בעל גודל קבוע, המתאים לקלט של רשת המדיניות של ה-RL.

האתגר המרכזי כאן הוא כיצד לעבד את מקטעי נתוני רמה 2, שעשויים להיות בעלי מבנה משתנה או גודל משתנה (תלוי באסטרטגיית החלוקה), לייצוג סטנדרטי. טכניקות אפשריות כוללות:

* **הנדסת מאפיינים מסכמת:** כפי שנדון בפרק 2.3.2, חילוץ סט קבוע של מאפיינים סטטיסטיים או מבוססי-דפוסים מכל מקטע Mi​.
* **שיטות הטבעה (Embedding):** שימוש ברשתות נוירונים (למשל, רשתות קונבולוציה או רשתות רקורנטיות) כדי ללמוד ייצוג דחוס ואינפורמטיבי של המקטע Mi​.
* **חלונות קבועים:** אם המקטעים הם זמניים, ניתן להשתמש בחלון זמן קבוע ולהשלים/לקטום נתונים לפי הצורך.

מחקר מציע הפשטת תצפיות על בסיס יחסים מרחביים לתוך ייצוג מבוסס-גרף (MARC), רעיון שאף על פי שהוא מיועד למשימות מרחביות, מדגיש את החשיבות של יצירת ייצוג מובנה ומופשט מהתצפית הגולמית. בספריות MARL כמו Ray RLlib, תצפיות מוחזרות כמילונים הממפים מזהי סוכנים לתצפיות האישיות שלהם , מה שמתאים למבנה המוצע.

### **3.4. מינוף ספריות ומסגרות MARL: Ray RLlib, PettingZoo, FinRL**

מימוש מערכת כה מורכבת יסתמך במידה רבה על ספריות קיימות. שלוש ספריות עיקריות הוזכרו והן רלוונטיות:

* **Ray RLlib:** ספריית קוד פתוח ללמידת חיזוק, התומכת בעומסי עבודה מבוזרים ובקנה מידה גדול, ומתאימה ליישומים תעשייתיים. היא מציעה תכונות MARL חשובות, כולל תמיכה במקרי שימוש מורכבים מרובי סוכנים (כגון משחק עצמי, שיתוף מידע, והפעלה עצמאית של סוכנים) ואלגוריתמים מודולריים. התיעוד של RLlib מפרט כיצד לבנות סביבות מרובות סוכנים מותאמות אישית, כולל הגדרת מזהי סוכנים ומילוני תצפיות/תגמולים/סיום, שהם קריטיים למימוש הסביבה הפיננסית המוצעת. RLlib משתלבת היטב עם כלי אופטימיזציית היפר-פרמטרים כמו Optuna דרך Ray Tune.
* **PettingZoo:** מהווה תקן API עבור סביבות MARL, בדומה ל-Gymnasium עבור RL של סוכן יחיד. היא תומכת בסביבות עם פעולות סדרתיות (AEC API) וסימולטניות (Parallel API), ומספקת מגוון סביבות ייחוס וכלי עזר. PettingZoo מאפשרת יצירת סביבות מותאמות אישית , וקיימים מדריכים לשילובה עם RLlib. סביר להניח שהסביבה הפיננסית המותאמת אישית תיבנה בהתאם לתקן של PettingZoo.
* **FinRL:** מסגרת קוד פתוח ייעודית ללמידת חיזוק פיננסית. היא בנויה בארכיטקטורה תלת-שכבתית (שכבת נתוני שוק, שכבת סוכנים, שכבת יישומים) וכוללת צנרת "אימון-בדיקה-מסחר". FinRL תומכת במגוון מקורות נתונים (אם כי לא בהכרח נתוני רמה 2 של רית'מיק באופן ישיר) ובסוכני DRL שונים, כולל אינטגרציה עם RLlib. FinRL-Meta שואפת להקל על בניית סביבות שוק ולהפחית את נטל עיבוד הנתונים המקדים. דוגמה רלוונטית היא סביבת MARL מותאמת אישית למכירה פומבית כפולה מתמשכת (CDA) המשתמשת ב-RLlib, בה הנתונים נוצרים ממשחק עצמי של הסוכנים, והתצפיות, הפעולות והתגמולים מוגדרים באופן ספציפי לסביבה הפיננסית. עקרונות ורכיבים מ-FinRL יכולים להנחות את תכנון הסביבה המותאמת אישית.

**שילוב:** התכנון הסביר ביותר יכלול יצירת סביבת מסחר מותאמת אישית העומדת בתקן PettingZoo. סביבה זו תהיה אחראית על קליטת נתוני רמה 2, עדכוני משטר מ-MMD, חלוקת הנתונים למטריצות Mi​, והגדרת האינטראקציה עם הסוכנים (תצפיות, פעולות, תגמולים). לאחר מכן, ניתן יהיה להשתמש ב-Ray RLlib כדי לאמן את סוכני ה-MARL (עם המדיניות המשותפת) בתוך סביבה מותאמת אישית זו.

הטבלה הבאה מסכמת את התאמתן של מסגרות MARL אלו למערכת המוצעת:

**טבלה 3: השוואת מסגרות MARL למסחר פיננסי**

| **מסגרת** | **תכונות MARL מרכזיות** | **סקלביליות** | **תמיכה בסביבה מותאמת אישית (קלות יצירת סביבות פיננסיות)** | **התאמה למסחר בחוזים עתידיים עם נתוני רמה 2** | **שילוב עם כלי אופטימיזציית היפר-פרמטרים (למשל, Optuna)** | **מקורות רלוונטיים** |
| --- | --- | --- | --- | --- | --- | --- |
| **Ray RLlib** | API גמיש למרובה סוכנים, תמיכה באימון מבוזר, אלגוריתמים מודולריים, CTDE. | גבוהה מאוד, מיועדת למערכות גדולות. | גבוהה; דורשת מימוש של MultiAgentEnv. התיעוד מפרט כיצד. | גבוהה; יכולה להתמודד עם מרחבי מצב/פעולה מורכבים הנדרשים. | כן, דרך Ray Tune. |  |
| **PettingZoo** | תקן API ל-MARL (AEC/Parallel), מגוון סביבות ייחוס, כלי עזר, עטיפות (wrappers). | תלוי בספריית ה-RL המשמשת לאימון (למשל, RLlib). | גבוהה; מספקת הדרכה ברורה ליצירת סביבות מותאמות אישית. | גבוהה; ה-API גמיש מספיק לייצוג סביבות פיננסיות מורכבות. | כן, כסביבה עבור ספריות RL שתומכות ב-Optuna. |  |
| **FinRL** | מסגרת ייעודית לפיננסים, מבנה שכבתי, צנרת אימון-בדיקה-מסחר, תמיכה במקורות נתונים פיננסיים. | תלוי בסוכן ה-DRL הנבחר (תומכת ב-RLlib). | בינונית-גבוהה; FinRL-Meta שואפת להקל על כך. דוגמת CDA מראה אפשרות ל-MARL מותאם. | בינונית-גבוהה; תומכת בנתוני שוק שונים, אך ייתכן שיידרש פיתוח נוסף לנתוני רמה 2 ספציפיים ולמבנה המטריצות. | כן, דרך ספריית ה-RL הבסיסית (למשל, RLlib עם Ray Tune). |  |

האתגר המרכזי בתכנון מערכת ה-MARL נובע מהשילוב הייחודי של מדיניות משותפת עם קלט נתונים אקראי ומבוזר. אם המדיניות המשותפת (הכוללת 4 סינרגיות) מורכבת, ומקטעי הנתונים Mi​ נבדלים מאוד זה מזה או דלים במידע, הסוכנים עלולים ליישם את המדיניות באופן שגוי או לא עקבי. הדבר עלול להוביל לפעולות תת-אופטימליות או אף סותרות, למרות הכוונה לשיתוף פעולה. "שיתוף הפעולה המושלם" במקרה זה עשוי שלא לנבוע מלמידה של הסוכנים כיצד לשתף פעולה, אלא מהיכולת של המערכת להבטיח שפעולותיהם, הנגזרות מאותה מדיניות אך מיושמות על נתונים שונים, יהיו קוהרנטיות ברמה הגלובלית. הדבר מצביע על כך שהצלחת המערכת תלויה באופן קריטי ביכולת של המדיניות המשותפת להתמודד עם מגוון רחב של תשומות ממקטעי הנתונים השונים, ובכך שהסכום של הפעולות הלוקאליות-אופטימליות (הנגזרות מהמדיניות המשותפת על נתונים מקומיים) אכן יתקרב לאסטרטגיה גלובלית-אופטימלית.

## **פרק 4: שילוב סופרפוזיציה בהשראת קוונטים לשיפור עמידות הסוכן**

הרעיון של שילוב "סופרפוזיציה" בהשראת מכניקת הקוונטים במערכת למידת חיזוק קלאסית הוא היבט חדשני ומאתגר של הפרויקט. המטרה היא לבחון כיצד עקרונות אלו יכולים לתרום להגברת העמידות (robustness) והאמינות (reliability) של סוכני המסחר. חשוב להדגיש כי הכוונה אינה לבנות מחשב קוונטי, אלא לשאוב השראה מרעיונות קוונטיים ולממשם באמצעות אלגוריתמים קלאסיים.

### **4.1. פענוח "סופרפוזיציה" בהקשר של סוכני RL**

#### **4.1.1. קישורים תיאורטיים לפיזיקה קוונטית (לפי הרצאת המשתמש)**

בפיזיקה קוונטית, סופרפוזיציה היא עיקרון יסודי לפיו מערכת קוונטית (כגון קיוביט) יכולה להתקיים במספר מצבים בו-זמנית עד לרגע המדידה, שבו המערכת "קורסת" לאחד המצבים האפשריים. לדוגמה, קיוביט יכול להיות בו-זמנית במצב '0' ובמצב '1', עם אמפליטודות הסתברות מסוימות לכל מצב. מחקרים בתחום למידת חיזוק קוונטית (QRL) מנצלים עיקרון זה, למשל, על ידי קידוד מצבים ופעולות מרובים במקביל באמצעות סופרפוזיציה, מה שמאפשר חקירה מקבילה של מרחב המצב-פעולה. סופרפוזיציה מאפשרת למערכות קוונטיות לעבד אפשרויות רבות בו-זמנית, מה שתורם ליעילות חישובית.

#### **4.1.2. אנלוגיות מעשיות: אלגוריתמים קלאסיים בהשראת קוונטים**

האתגר הוא לתרגם את הרעיון המופשט של סופרפוזיציה למושגים מעשיים הניתנים למימוש במערכת RL קלאסית. מספר גישות "בהשראת קוונטים" יכולות לשמש כאנלוגיה:

* **חקירה מקבילה של מצבים/מדיניויות:** ניתן לדמות את יכולת החקירה המקבילה של QRL על ידי שימוש באנסמבל של מדיניויות או על ידי שמירה על התפלגות הסתברות על פני פעולות אפשריות.
* **ייצוגים הסתברותיים:** סוכן יכול לשמור על "אמונה" הסתברותית לגבי מספר מצבי שוק אפשריים או מספר פעולות אופטימליות, במקום להתחייב למצב או פעולה בודדים.
* **אלגוריתמים קלאסיים בהשראת קוונטים:** מחקרים מציעים אלגוריתמים קלאסיים המנצלים עקרונות קוונטיים כמו סופרפוזיציה. לדוגמה, QIDDM (Quantum-Inspired Dynamic Decision-Making Algorithm) הוא אלגוריתם קלאסי שבו החלטות קיימות כסופרפוזיציה הסתברותית עד שתצפיות חיצוניות גורמות ל"קריסה" להחלטה דטרמיניסטית. מסגרת Q-Policy מקודדת פונקציות ערך בסופרפוזיציה קוונטית לצורך הערכה מקבילה של זוגות מצב-פעולה, ומאומתת באמצעות אמולציה קלאסית.

### **4.2. כיצד סופרפוזיציה יכולה להגביר עמידות ואמינות במערכות מסחר MARL**

שילוב טכניקות בהשראת סופרפוזיציה יכול לתרום לעמידות ואמינות המערכת במספר דרכים:

* **עמידות הנובעת מגיוון (Robustness from Diversity):** אם "סופרפוזיציה" מתורגמת לכך שסוכן שוקל או מחזיק מספר השערות (hypotheses) בו-זמנית לגבי מצב השוק או הפעולות האופטימליות, הוא יכול להיות עמיד יותר לרעש או לשינויים פתאומיים בשוק. לדוגמה:  
  + **אנסמבל של מדיניויות:** אם מדיניות אחת באנסמבל נכשלת בתנאי שוק מסוימים, מדיניויות אחרות עשויות להצליח, מה שמוביל לעמידות כוללת של המערכת.
  + **חקירה מגוונת:** סוכן הנמצא ב"סופרפוזיציה" של אסטרטגיות פוטנציאליות יכול לחקור באופן אינהרנטי מגוון רחב יותר של תגובות, מה שמגביר את יכולתו להתמודד עם תרחישים בלתי צפויים.
  + **התמודדות עם אי-ודאות:** טכניקות בהשראת קוונטים יכולות לסייע בבניית מערכות למידת מכונה עמידות ומסתגלות יותר על ידי ניצול עקרונות של אי-ודאות וסופרפוזיציה. אלגוריתם QIDDM, למשל, הראה שיפור בדיוק ההחלטות והפחתת התחייבויות מוקדמות בסביבות לא ודאיות, כולל סביבות פיננסיות.
* **אמינות הנובעת מהתחייבות מאוחרת/קונצנזוס (Reliability from Delayed Commitment/Consensus):**
  + אם סוכן נמצא ב"סופרפוזיציה" של פעולות פוטנציאליות, הוא עשוי "לקרוס" לפעולה סופית רק לאחר איסוף ראיות נוספות או השגת קונצנזוס (למשל, עם סוכנים אחרים או על סמך אותות חזקים יותר מהשוק). התחייבות מאוחרת זו יכולה להוביל להחלטות אמינות יותר.
  + אלגוריתם QIDDM, לדוגמה, דוחה התחייבויות מוקדמות ומאפשר תגובות מסתגלות. ההחלטה מתקבלת רק כאשר הפחתת האנטרופיה (כלומר, הגידול במידע) עוברת סף מסוים.

היתרון המרכזי של "סופרפוזיציה" בהקשר זה אינו טמון בחישוב קוונטי אמיתי, אלא ביכולת לאפשר למערכת ה-RL לנהל אי-ודאות וגיוון בצורה יעילה יותר. סוכן, או קבוצת סוכנים, הפועלים כישות "בסופרפוזיציה" (למשל, באמצעות אנסמבלים), יכולים לחקור אסטרטגיות מגוונות בו-זמנית. כישלון של אסטרטגיה אחת בתנאי שוק נתונים עשוי להיות מקוזז על ידי הצלחה של אסטרטגיה אחרת, מה שמוביל לעמידות מערכתית כוללת. "קריסת" המערכת להחלטה ספציפית יכולה להיות מנגנון אגרגציה או בחירה מושכלת.

### **4.3. אסטרטגיות אינטגרציה עבור מערכת המסחר MARL**

קיימות מספר אסטרטגיות קלאסיות למימוש רעיון הסופרפוזיציה במערכת ה-MARL:

1. **שיטות אנסמבל (Ensemble Methods):** ניתן לאמן מספר מדיניויות (או גרסאות של המדיניות המשותפת עם שינויים קלים או אתחולים שונים) ולבצע אגרגציה של פעולותיהן (למשל, באמצעות הצבעה, ממוצע משוקלל של הסתברויות פעולה). כל מדיניות באנסמבל יכולה לייצג "מצב" בסופרפוזיציה. מחקרים מראים כי למידת חיזוק באנסמבל (ERL) מנצלת מספר מודלים לחקירה מקיפה של מרחב הבעיה ומפגינה יכולות הכללה חזקות. ניתן ליישם אנסמבלים על ידי הצבעה על פעולות הסוכנים או ממוצע משוקלל של הסתברויות הפעולה, כאשר המשקולות יכולות להתבסס על מדדי ביצועים כמו יחס שארפ.
2. **מדיניויות הסתברותיות (Probabilistic Policies):** במקום שהמדיניות תפיק פעולה דטרמיניסטית, היא יכולה להפיק התפלגות הסתברות על פני מרחב הפעולות. דגימה מהתפלגות זו או שימוש בפרמטריה יכולים לייצג את המצב ה"מורכב" (superposed).
3. **למידת חיזוק מרובת-השערות / מודלים של מצב חבוי (Multi-Hypothesis RL / Latent State Models):** הסוכן יכול לשמור על אמונות (beliefs) לגבי מספר מצבי שוק חבויים אפשריים, ובכך להיות למעשה ב"סופרפוזיציה" של אמונות אלו, ולפעול על סמך התפלגות זו. מחקר מציע שיטת MARL (NDD) שבה התגמול הגלובלי הרועש מקורב על ידי מודל תערובת גאוסיאנית (GMM), והסוכנים לומדים פונקציות ערך התפלגותיות מקומיות. למידת חיזוק התפלגותית (Distributional RL) ממדלת את ערך המצב-פעולה כהתפלגות, ולא כערך יחיד.
4. **חקירה מקבילה של מרחב המדיניות/מצב (Parallel Exploration of Policy/State Space):** שימוש בטכניקות המאפשרות לסוכן או למערכת לחקור חלקים שונים של מרחב המדיניות או המצב באופן מקביל, בדומה למקביליות קוונטית.
5. **מנגנון דמוי-QIDDM (QIDDM-like Mechanism):** ניתן לממש תהליך קבלת החלטות קלאסי שבו פעולות פוטנציאליות (הנגזרות מארבע הסינרגיות) מקבלות הסתברויות (או "אמפליטודות"). הסוכן נשאר ב"סופרפוזיציה" זו עד ש"טריגר קריסה" (למשל, אות בעל ביטחון גבוה מהנתונים, קונצנזוס בין תת-קבוצה של סוכנים, שינוי משטר ספציפי) מאלץ בחירה בפעולה ספציפית. גישה זו מתאימה ישירות לרעיון המקורי של המשתמש.

שילוב רעיונות בהשראת סופרפוזיציה, למשל באמצעות אנסמבלים, משנה באופן מהותי את הדינמיקה של מערכת ה-MARL. במקום N סוכנים שלכל אחד מדיניות אחת, ייתכן שיהיו N סוכנים שכל אחד מהם מריץ *פנימית* אנסמבל של M מדיניויות, או ש-N הסוכנים עצמם מהווים אנסמבל. הדבר מגדיל את המורכבות החישובית אך עשוי לשפר משמעותית את העמידות לאי-סטציונריות ולרעש, שהם נפוצים בשווקים פיננסיים. הערכת מערכת כזו תדרוש מדדים הלוכדים לא רק רווח והפסד, אלא גם את גיוון החקירה ואת יכולת המערכת להסתגל בצורה חיננית לזעזועי שוק בלתי צפויים. ה"סופרפוזיציה" הופכת למנגנון לחקירה וניצול מובנים של אסטרטגיות מגוונות.

הטבלה הבאה מפרטת גישות אפשריות לשילוב סופרפוזיציה בהשראת קוונטים:

**טבלה 4: גישות סופרפוזיציה בהשראת קוונטים לעמידות סוכני RL**

| **גישה** | **רעיון "בהשראת קוונטים" מרכזי** | **מנגנון לעמידות/אמינות** | **תרשים מימוש מעשי למסחר MARL** | **מקורות רלוונטיים** |
| --- | --- | --- | --- | --- |
| ביצוע מדיניות באנסמבל | חקירה מקבילה של "מצבי" מדיניות מרובים. | גיוון אסטרטגי; אם מדיניות אחת נכשלת, אחרות עשויות להצליח. אגרגציית החלטות (למשל, הצבעה) יכולה לסנן רעשים. | כל אחד מ-N הסוכנים (או המערכת כולה) מריץ K גרסאות של המדיניות המשותפת (למשל, עם פרמטרים שונים במקצת, או מאומנות על תתי-דגימות שונות של נתונים). החלטת המסחר הסופית מתקבלת על ידי הצבעת רוב או ממוצע משוקלל של המלצות ה-K מדיניויות. |  |
| בחירת פעולה הסתברותית | ייצוג מצב קוונטי כהתפלגות על פני מצבי בסיס. | הימנעות מהתחייבות מוקדמת לפעולה בודדת; מאפשר תגובה גמישה יותר לאי-ודאות. | רשת המדיניות המשותפת מפיקה התפלגות הסתברות על פני מרחב הפעולות האפשרי (למשל, קנה, מכור, החזק, עם הסתברויות שונות). הפעולה הסופית נדגמת מהתפלגות זו, או שנבחרת הפעולה בעלת ההסתברות הגבוהה ביותר לאחר שעברה סף ביטחון מסוים. |  |
| מעקב אחר השערות מצב מרובות | סוכן "קיים" במספר מצבי אמונה לגבי הסביבה. | יכולת להגיב למגוון רחב יותר של התפתחויות שוק אפשריות; פחות רגישות לזיהוי שגוי של המצב הנוכחי. | כל סוכן מחזיק התפלגות אמונה על פני מספר מצבי שוק חבויים (למשל, "מגמה עולה חזקה", "דשדוש תנודתי"). המדיניות המשותפת מותנית על התפלגות אמונה זו, או שכל השערת מצב מזינה עותק של המדיניות, וההחלטות משוקללות. |  |
| התחייבות מאוחרת דמוית-QIDDM | אנלוגיה ל"קריסת פונקציית הגל" הקוונטית. | החלטות מתקבלות רק כאשר יש מספיק "ראיות" או ביטחון, מה שמפחית פעולות פזיזות בסביבות רועשות. | הסוכנים מעריכים במקביל את התועלת של הפעלת כל אחת מארבע הסינרגיות. הם נשארים ב"סופרפוזיציה" של סינרגיות פוטנציאליות. "קריסה" לסינרגיה ספציפית (ולפעולות הנגזרות ממנה) מתרחשת כאשר תנאי שוק מסוימים (למשל, אות חזק ממקטע הנתונים Mi​, אישור ממשטר ה-MMD, או קונצנזוס בין סוכנים) עוברים סף קריטי. |  |

תרגום רעיון ה"סופרפוזיציה" לטכניקות RL קלאסיות הניתנות למימוש הוא צעד חיוני. טבלה זו מציעה מגוון אפשרויות מעשיות להשגת העמידות והאמינות שהמשתמש שואף אליהן, ומגשרת בין ההשראה הפיזיקלית לפתרון ההנדסי.

## **פרק 5: אימון, הערכה ואופטימיזציה של ביצועים**

לאחר תכנון ארכיטקטורת ה-MARL ושילוב הרעיונות המתקדמים, השלב הבא הוא אימון המערכת, הערכת ביצועיה באופן קפדני, ואופטימיזציה של הפרמטרים השונים. שלב זה חיוני להבטחת יעילותה ועמידותה של המערכת בסביבת השוק הדינמית והרועשת.

### **5.1. הגדרת קריטריונים להצלחה ויעדי רווח והפסד**

הגדרת קריטריונים ברורים להצלחה היא צעד ראשון והכרחי. בעוד שהמשתמש ציין יעד ספציפי של "כמה האסטרטגיה צריכה להרוויח בשבועיים", חשוב להבין כי התמקדות צרה ביעד רווח והפסד (P&L) קצר טווח עלולה להוביל להתאמת יתר (overfitting). לכן, יש לאמץ מגוון רחב יותר של מדדי ביצועים:

* **מדדי תשואה מותאמת לסיכון:** יחס שארפ (Sharpe Ratio), יחס סורטינו (Sortino Ratio), יחס קלמאר (Calmar Ratio). מדדים אלו מעריכים את התשואה ביחס לרמת הסיכון שנלקחה.
* **מדדי רווחיות:** תשואה מצטברת (Cumulative Return), פקטור רווח (Profit Factor – סך הרווח חלקי סך ההפסד).
* **מדדי סיכון:** משיכה מקסימלית (Maximum Drawdown – MDD), תנודתיות התשואות.
* **מדדי יעילות מסחר:** אחוז עסקאות רווחיות (Win Rate), תוחלת רווח לעסקה (Expectancy).

הצלחת המערכת צריכה להימדד גם ביכולתה לשמור על ביצועים עקביים וחזקים על פני משטרי שוק שונים (כפי שזוהו על ידי מנוע ה-MMD) ועל פני מכשירים פיננסיים שונים. היעד של "רווח תוך שבועיים" יכול לשמש כנקודת ייחוס להערכה, אך אין לבצע אופטימיזציה ישירה ובלעדית אליו.

### **5.2. עיצוב פונקציית תגמול עבור MARL שיתופי**

פונקציית התגמול (R) היא קריטית בלמידת חיזוק, שכן היא מנחה את תהליך הלמידה של הסוכנים. במערכת MARL שיתופית, פונקציית התגמול צריכה לשקף את מטרות הצוות הכוללות.

#### **5.2.1. איזון בין יעדי רווח והפסד לניהול סיכונים והימנעות מהתאמת יתר**

פונקציית תגמול המבוססת אך ורק על P&L היא מסוכנת ונוטה להתאמת יתר. יש לשלב בה רכיבים המעודדים ניהול סיכונים ומפחיתים את הסיכוי ללמידת אסטרטגיות שבריריות.

* **שילוב מדדי סיכון:** ניתן לשלב ישירות מדדים כמו יחס שארפ או סורטינו בפונקציית התגמול.
* **עונשים על סיכון:** להטיל עונשים על משיכות גדולות, תנודתיות גבוהה של התיק, או חריגה ממגבלות סיכון מוגדרות מראש.
* **תגמול על עמידות (Robustness Rewards):** ניתן לשקול תגמול על ביצועים עקביים על פני חלונות זמן מתגלגלים או משטרי שוק שונים.
* **מניעת "ריגול אחר התגמול" (Reward Hacking):** יש לוודא שהסוכנים אינם לומדים "לרמות" את פונקציית התגמול על ידי ניצול פרצות או התנהגויות לא רצויות שמובילות לתגמול גבוה באופן מלאכותי. שימוש בטכניקות של עיצוב תגמול (reward shaping) או למידת חיקוי (imitation learning) יכול לסייע בהפחתת רעש באות התגמול ולהפוך אותו לחזק יותר.

#### **5.2.2. שילוב ארבע הסינרגיות במרחב התגמול/פעולה**

פונקציית התגמול צריכה לעודד את הסוכנים לבצע את ארבע הסינרגיות המוגדרות באסטרטגיה המשותפת, כאשר התנאים מתאימים. ניתן לעשות זאת על ידי:

* **תגמול ספציפי לסינרגיה:** מתן תגמול נוסף כאשר הסוכן מזהה ומבצע בהצלחה סינרגיה מסוימת התואמת את מצב השוק (כפי שהוא נתפס דרך Mi​ ופרמטרי המשטר).
* **עיצוב תגמול מבוסס-מטרה:** אם כל סינרגיה קשורה למטרה ספציפית (למשל, לכידת מגמה, ניצול תנודתיות), ניתן לתגמל את הסוכן על השגת מטרות אלו. מחקר מתאר תהליך בו התנהגות תיאום מתוארת בשפה טבעית ומתורגמת לפונקציית תגמול, רעיון שניתן להתאים לתרגום הסינרגיות לרכיבי תגמול.

#### **5.2.3. הקצאת אשראי לתרומות סוכנים בודדים ליעדים משותפים**

במערכת MARL שיתופית עם תגמול גלובלי, אתגר מרכזי הוא בעיית הקצאת האשראי (credit assignment): כיצד לקבוע את תרומתו של כל סוכן בודד להצלחה או לכישלון הכולל של הצוות? אם סוכן j ביצע פעולה על סמך Mj​ שהובילה להפסד כולל, האם זה בגלל פעולתו, פעולת סוכן אחר, או שהמדיניות המשותפת פשוט לא התאימה לתנאי השוק באותו רגע?

* **Value Decomposition Networks (VDN):** שיטות כמו VDN מנסות לפרק את פונקציית הערך הגלובלית לסכום של פונקציות ערך אינדיבידואליות, ובכך להעריך את תרומתו של כל סוכן.
* **Counterfactual Rewards:** הערכת תרומת סוכן על ידי השוואת התגמול שהתקבל עם התגמול שהיה מתקבל אילו הסוכן היה פועל אחרת.
* **Shapley Values:** שימוש בערכי שפלי לחלוקת התגמול בין הסוכנים בהתאם לתרומתם השולית.
* **Influence Scope of Agents (ISA):** אלגוריתם המחשב את היקף ההשפעה של סוכנים על מצבים, ומשתמש בתלות ההדדית בין פעולות סוכנים לתכונות מצב לצורך הקצאת אשראי.
* **Episodic Control for Credit Assignment:** שימוש בזיכרון אפיזודי לאחסון וקיבוץ מסלולים קודמים, ותגמול פנימי לשיפור הערכת ערך אישית ומשותפת.

הטבלה הבאה מציגה רכיבים אפשריים לפונקציית תגמול חזקה עבור מערכת המסחר MARL השיתופית:

**טבלה 5: רכיבי פונקציית תגמול חזקה למסחר MARL שיתופי**

| **רכיב תגמול** | **רעיון/מדד ניסוח** | **רציונל להכללה** | **משקל/חשיבות פוטנציאליים** | **מקורות רלוונטיים** |
| --- | --- | --- | --- | --- |
| רווח והפסד (P&L) של התיק (מבוסס צוות) | ΔPortfolioValuet​ | מדד ביצועים מרכזי, משקף את המטרה העיקרית של מסחר. | גבוה, אך מאוזן עם רכיבי סיכון. |  |
| תשואה מותאמת לסיכון | יחס שארפ, יחס סורטינו, מחושבים על חלון מתגלגל. | מעודד רווחיות תוך ניהול סיכונים. | גבוה. |  |
| ציון ביצוע סינרגיה (לכל אחת מ-4 הסינרגיות) | +1 אם סינרגיה k זוהתה ובוצעה בהצלחה בתנאים מתאימים, 0 אחרת. או תגמול רציף המבוסס על איכות ביצוע הסינרגיה. | מבטיח שהסוכנים לומדים לבצע את כל רכיבי האסטרטגיה המשותפת. | בינוני-גבוה, תלוי בחשיבות כל סינרגיה. | (בהשראה) |
| מדד שיתוף פעולה/קוהרנטיות | עונש על פעולות סותרות בין סוכנים (אם ניתן לזיהוי), או תגמול על פעולות מתואמות התומכות במטרה גלובלית. | חיוני להשגת "שיתוף פעולה מושלם". | בינוני. |  |
| רכיב נגד התאמת יתר | עונש על מורכבות מדיניות, תגמול על גיוון בפעולות (אם רצוי במסגרת הסופרפוזיציה), או שימוש בתגמול מבוסס-חיקוי. | מפחית סיכון ללמידת אסטרטגיות שבריריות. | בינוני. |  |
| עונשי עלויות | עמלות עסקה, החלקה (slippage), עלויות החזקת פוזיציה. | מבטיח שהאסטרטגיה רווחית גם לאחר התחשבות בעלויות מסחר ריאליות. | גבוה. |  |

### **5.3. אופטימיזציה של היפר-פרמטרים (למשל, שימוש ב-Optuna עם Ray)**

מערכת MARL מורכבת כוללת מספר רב של היפר-פרמטרים שיש לכייל, כגון שיעורי למידה, ארכיטקטורות רשת, מקדם היוון (discount factor), פרמטרי חקירה (exploration parameters), מספר הסוכנים/מחיצות הנתונים, ופרמטרי מנוע ה-MMD. אופטימיזציה ידנית של פרמטרים אלו אינה מעשית.

ספריות כמו Optuna ו-Ray Tune (המשולבת עם RLlib) מספקות כלים לאוטומציה של תהליך זה. Optuna משתמשת באופטימיזציה בייסיאנית ובאלגוריתמים מתקדמים אחרים לחיפוש יעיל במרחבי היפר-פרמטרים גדולים. התהליך כולל הגדרת פונקציית "מטרה" (objective function) העוטפת את תהליך אימון המודל (במקרה זה, אימון מערכת ה-MARL על סט אימון והערכתה על סט ולידציה) ומחזירה מדד ביצועים (למשל, יחס שארפ על סט הולידציה) שאותו Optuna מנסה למקסם או למזער. Ray Tune מאפשרת לבצע חיפושי היפר-פרמטרים אלו באופן מבוזר ומקבילי, מה שמאיץ משמעותית את התהליך. שילוב של Optuna עם Ray Tune (דרך OptunaSearch) הוא גישה חזקה במיוחד.

### **5.4. מסגרת בדיקות לאחור (Backtesting) עבור מערכות מסחר MARL**

בדיקות לאחור (backtesting) הן כלי חיוני להערכת אסטרטגיות מסחר, אך הן חייבות להתבצע בקפדנות, במיוחד עבור מערכות MARL הפועלות בשווקים לא-סטציונריים כמו שוק החוזים העתידיים.

#### **5.4.1. טיפול באי-סטציונריות ושינויי משטר**

שווקים פיננסיים הם דינמיים ומשתנים תדיר (לא-סטציונריים). מנוע זיהוי המשטירים באמצעות MMD נועד להתמודד ישירות עם היבט זה. מסגרת הבדיקות לאחור צריכה:

* **להשתמש בנתונים היסטוריים איכותיים ומקיפים:** נתוני רמה 2, כולל עומק ספר פקודות, חיוניים. עבור חוזים עתידיים, יש לטפל בגלגול חוזים (contract rollover).
* **לבצע ולידציה חוצת-משטרים:** להעריך את ביצועי המערכת על פני תקופות המייצגות משטרי שוק שונים, כפי שזוהו על ידי מנוע ה-MMD.
* **להשתמש בטכניקות בדיקה מתקדמות:**
  + **Walk-Forward Analysis:** מחלקים את הנתונים ההיסטוריים לתקופות אימון (in-sample) ובדיקה (out-of-sample) מתגלגלות. המערכת מאומנת על תקופת אימון ונבדקת על תקופת הבדיקה הבאה אחריה, והתהליך חוזר על עצמו. זה מדמה טוב יותר תנאי מסחר אמיתיים ומסייע בהפחתת התאמת יתר.
  + **Monte Carlo Simulations & Stress Testing:** לבחינת עמידות האסטרטגיה בתרחישים קיצוניים או "ברבורים שחורים".
  + **Time-Series Cross-Validation:** שיטות ולידציה צולבת המותאמות לסדרות עתיות (למשל, blocked cross-validation).
* **למדל עלויות מסחר ריאליסטיות:** עמלות, החלקה (slippage), והשפעה על השוק (market impact).
* **להימנע מהטיות נפוצות:** הטיית התאמת יתר (overfitting bias), הטיית מבט לעתיד (look-ahead bias), והטיית שרידות (survivorship bias).

מחקר מציע למדל את המסחר כבעיית RL לא-סטציונרית ולאפשר בחירה דינמית של סוכן ה-RL הטוב ביותר על סמך ביצועי רווח, במקום להסתמך על משטרים קודמים.

#### **5.4.2. בדיקה על מסגרות זמן ומכשירים שונים**

כפי שביקש המשתמש, יש לבדוק כיצד המערכת מתפקדת על פני מסגרות זמן שונות (למשל, תוך-יומי, יומי) ועל פני מכשירים פיננסיים שונים (חוזים עתידיים על מדדים, סחורות, מט"ח). בדיקה זו חיונית להערכת יכולת ההכללה (generalization) של המערכת. ייתכן שיידרש אימון מחדש או כוונון עדין (fine-tuning) של המערכת עבור כל מכשיר או מסגרת זמן, שכן מחקרים מצביעים על כך שביצועי אסטרטגיות עשויים להיות ספציפיים לנכס.

### **5.5. התייחסות לחששות מהתאמת יתר, במיוחד עם אופטימיזציה של רווח והפסד**

אופטימיזציה ישירה של יעד P&L קצר טווח (למשל, "להרוויח X בשבועיים") היא מתכון כמעט בטוח להתאמת יתר. כדי למתן סיכון זה:

* **פונקציות תגמול חזקות:** שימוש במדדים מותאמי סיכון (ראה פרק 5.2.1) במקום P&L גולמי.
* **רגולריזציה:** הוספת רכיבי רגולריזציה לפונקציית המטרה של אלגוריתם ה-RL (למשל, L1/L2 על משקולות הרשת, עונש על מורכבות המדיניות).
* **ולידציה קפדנית:** שימוש נרחב בבדיקות walk-forward ובנתוני out-of-sample.
* **בדיקות הכללה:** הערכת ביצועים על מכשירים ומסגרות זמן שונות.
* **עמידות בהשראת סופרפוזיציה:** אם רעיון הסופרפוזיציה (פרק 4) ממומש באמצעות גיוון מדיניויות או אנסמבלים, הדבר עשוי לתרום להפחתת התאמת יתר על ידי מניעת התמקדות יתר באסטרטגיה בודדת ושברירית.
* **מניעת "ריגול אחר התגמול":** עיצוב זהיר של פונקציית התגמול כדי למנוע מהסוכנים ללמוד קיצורי דרך לא רצויים. שילוב למידת חיקוי עם RL יכול להפוך את אות התגמול לחזק יותר לרעש, ובכך להפחית התאמת יתר לאותות P&L מזויפים.

המתח בין הרצון להשיג יעד P&L ספציפי לבין הצורך בעמידות והימנעות מהתאמת יתר הוא מרכזי. מנוע זיהוי המשטרים באמצעות MMD והרעיון של "סופרפוזיציה" הם אמצעי נגד קריטיים: MMD מספק הקשר להתאמת האסטרטגיה, וסופרפוזיציה (אם ממומשת כגיוון מדיניות) מספקת חלופות אם אסטרטגיה אחת מתחילה להיכשל. לכן, פונקציית התגמול חייבת להיות מאוזנת בקפידה, ולכלול לא רק P&L אלא גם מדדי סיכון, ואולי אף רכיבים המעודדים ביצוע מגוון של הסינרגיות.

הטבלה הבאה מסכמת נקודות מרכזיות לבדיקות לאחור של מערכת המסחר המוצעת:

**טבלה 6: רשימת תיוג לפרוטוקול בדיקות לאחור עבור מערכת מסחר MARL בחוזים עתידיים**

| **נקודת בדיקה** | **שיקול/נוהג מומלץ** | **מקורות רלוונטיים** |
| --- | --- | --- |
| מקור ואיכות נתונים | נתוני רמה 2 (רית'מיק), טיפול בנתונים חסרים/שגויים, התחשבות במאפייני ספר פקודות (למשל, פקודות מוסתרות). |  |
| טיפול באי-סטציונריות | בדיקות Walk-Forward, ניתוח ספציפי למשטר (על בסיס פלט MMD), הערכת ביצועים על פני תקופות שוק שונות (שורי, דבי, תנודתי). |  |
| היבטים ספציפיים ל-MARL | יציבות מדיניות בין סוכנים, הערכת איכות שיתוף הפעולה, השפעת חלוקת הנתונים על ביצועי סוכנים בודדים ועל המערכת כולה. |  |
| היבטים ספציפיים לחוזים עתידיים | טיפול בגלגול חוזים, דרישות מינוף ובטחונות (margin), החלקה (slippage) ספציפית לשוק החוזים. |  |
| מידול עלויות | עמלות עסקה, עלויות החלקה ריאליסטיות, הערכת השפעה פוטנציאלית על השוק (במיוחד אם המערכת סוחרת בנפחים גדולים). |  |
| זיהוי התאמת יתר | שימוש קפדני בנתוני Out-of-Sample, מבחנים סטטיסטיים (למשל, בדיקת עמידות פרמטרים), השוואה לביצועי בנצ'מרק נאיביים. |  |
| מדדי ביצועים | P&L, מדדים מותאמי סיכון (שארפ, סורטינו), משיכה מקסימלית, פקטור רווח, שיעור הצלחה, הערכת ביצוע כל אחת מ-4 הסינרגיות. |  |
| הערכת "סופרפוזיציה קוונטית" | מדדי עמידות (למשל, ביצועים בתנאי רעש גבוהים, יציבות בתקופות זעזועים), גיוון בפעולות (אם רלוונטי למימוש הסופרפוזיציה). | פרק 4, |

הצלחה אמיתית של מערכת כה מורכבת אינה נמדדת רק בהשגת יעד P&L חד-פעמי, אלא ביכולתה המתמשכת ללמוד, להסתגל, ולשמור על רווחיות ועמידות בסביבת שוק המשתנה ללא הרף. הדבר מצביע על צורך ברמה מטא-אנליטית של הערכה: באיזו מהירות המערכת יכולה להתאים עצמה או ללמוד מחדש כאשר היא נתקלת בנתונים חדשים או במשטר שוק שלא נצפה באימונים? הדרישה לבדוק את המערכת על פני מסגרות זמן ומכשירים שונים היא צעד בכיוון זה, אך התכנון הכולל צריך לשאוף ליכולת למידה והסתגלות מתמשכת.

## **פרק 6: מפת דרכים ליישום ושיקולים מתקדמים**

מימוש מערכת המסחר מרובת הסוכנים כפי שתוארה הוא משימה מורכבת הדורשת תכנון קפדני וגישה שלבית. פרק זה יתווה מפת דרכים אפשרית ליישום וידון בשיקולים מתקדמים הקשורים לאתגרים המרכזיים.

### **6.1. תוכנית יישום שלב אחר שלב**

ניתן לחלק את תהליך הפיתוח והיישום למספר שלבים מרכזיים:

1. **שלב 1: הקמת תשתית נתונים:**
   * פיתוח צנרת נתונים אמינה לרכישה, אחסון ועיבוד של נתוני רמה 2 מרית'מיק בזמן אמת ובאופן היסטורי.
   * הבטחת איכות נתונים, טיפול בערכים חסרים, וסנכרון חותמות זמן.
2. **שלב 2: פיתוח ובדיקת מנוע זיהוי משטרי שוק (MMD):**
   * מימוש אלגוריתם MMD מסדר גבוה, תוך התבססות על המחקר של Horvath ו-Issa והקוד הזמין (למשל, issaz/signature-regime-detection).
   * בדיקה וכיול של מנוע ה-MMD על נתונים היסטוריים כדי לוודא את יכולתו לזהות שינויי משטר באופן משמעותי.
   * הגדרת אופן שילוב פרמטרי המשטר בזרם הנתונים הראשי.
3. **שלב 3: פיתוח בסיס ייחוס (Baseline) עם סוכן RL יחיד:**
   * מימוש אסטרטגיית המסחר המשותפת (כולל ארבע הסינרגיות) באמצעות סוכן RL יחיד הפועל על מלוא נתוני רמה 2 (לאחר הוספת פרמטרי המשטר, אך ללא חלוקה למטריצות).
   * שלב זה ישמש כנקודת השוואה להערכת התועלת של גישת ה-MARL.
4. **שלב 4: בניית סביבת MARL מותאמת אישית:**
   * פיתוח סביבת MARL מותאמת אישית, רצוי בהתאם לתקן PettingZoo.
   * הסביבה תהיה אחראית על:
     + קבלת נתוני רמה 2 מועשרים בפרמטרי משטר.
     + ביצוע החלוקה האקראית של הנתונים למטריצות Ma​,Mb​,...,Mn​.
     + העברת התצפיות המתאימות (מטריצה Mi​ ופרמטרי משטר) לכל סוכן i.
     + קבלת פעולות מהסוכנים וביצוען בסביבת המסחר המדומה.
     + חישוב והחזרת תגמולים (על בסיס פונקציית התגמול המשותפת).
   * ניתן להיעזר בדוגמאות כמו סביבת ה-CDA ב-FinRL ובהנחיות של RLlib ליצירת סביבות מרובות סוכנים.
5. **שלב 5: אימון סוכני MARL:**
   * אימון סוכני ה-MARL באמצעות המדיניות המשותפת ופונקציית התגמול השיתופית, תוך שימוש בספריית RLlib.
   * התמקדות בהשגת תיאום ושיתוף פעולה בין הסוכנים.
6. **שלב 6: שילוב שכבת סופרפוזיציה בהשראת קוונטים:**
   * מימוש אחת מהטכניקות שנדונו בפרק 4 (למשל, אנסמבל של מדיניויות, מדיניות הסתברותית, מנגנון דמוי-QIDDM) כדי לשפר את עמידות הסוכנים.
   * שילוב שכבה זו באופן שישתלב עם ארכיטקטורת ה-MARL הקיימת.
7. **שלב 7: בדיקות לאחור ואופטימיזציה:**
   * ביצוע בדיקות לאחור קפדניות על פי הפרוטוקול שנקבע (פרק 5.4).
   * שימוש ב-Optuna ו-Ray Tune לאופטימיזציה של היפר-פרמטרים של מערכת ה-MARL, מנוע ה-MMD, ושכבת הסופרפוזיציה.
8. **שלב 8: איטרציה ושיפור מתמיד:**
   * ניתוח תוצאות הבדיקות, זיהוי נקודות תורפה, וביצוע איטרציות לשיפור כל אחד מרכיבי המערכת.

### **6.2. התמודדות עם אתגרי מפתח: מדרגיות, אי-סטציונריות ונצפות חלקית ב-MARL**

כפי שצוין לאורך הדוח, MARL מציבה אתגרים ייחודיים :

* **מדרגיות (Scalability):** השימוש ב-Ray RLlib, המיועדת למערכות מבוזרות , הוא צעד בכיוון הנכון. חלוקת הנתונים למטריצות קטנות יותר עשויה גם היא לתרום למדרגיות, בתנאי שתקורה התקשורת (אם נדרשת תקשורת מפורשת בין סוכנים) מנוהלת היטב.
* **אי-סטציונריות (Non-stationarity):** מנוע זיהוי המשטירים באמצעות MMD הוא ניסיון ישיר להתמודד עם אי-סטציונריות השוק. רעיון הסופרפוזיציה, אם הוא מיושם כגיוון מדיניות, יכול גם הוא לסייע על ידי מתן אפשרות למערכת להסתגל טוב יותר לשינויים. טכניקות של אופטימיזציה בשיטת walk-forward ולמידה מתמשכת (continuous learning) הן חיוניות להתמודדות ארוכת טווח עם אי-סטציונריות.
* **נצפות חלקית (Partial Observability):** כל סוכן במערכת המוצעת פועל, מעצם הגדרתו, תחת נצפות חלקית (הוא רואה רק את מטריצת הנתונים Mi​ שלו). המדיניות המשותפת ופרמטרי המשטר הגלובליים נועדו לספק הקשר מספיק לפעולה קוהרנטית. אם יתברר שהתצפיות המקומיות אינן מספקות, ניתן יהיה לשקול הוספת מנגנוני תקשורת מפורשים בין הסוכנים כדי לאפשר להם לשתף מידע קריטי.

### **6.3. אגרגציה ומיזוג מדיניות מתשומות/תצפיות הטרוגניות של סוכנים**

אף על פי שכל הסוכנים חולקים את אותה מדיניות, הם מקבלים תשומות שונות (מטריצות Mi​ שונות). אם הפלט של כל סוכן (למשל, הסתברויות לפעולה) צריך לעבור אגרגציה להחלטת צוות אחת (למשל, לצורך ניהול סיכונים כולל של התיק, או לביצוע פקודה נטו אחת בשוק), יש לדון כיצד לבצע אגרגציה זו.

* **הצבעת רוב או ממוצע משוקלל:** ניתן לאסוף את המלצות הפעולה מכל הסוכנים ולבחור את הפעולה שקיבלה את רוב הקולות, או לחשב ממוצע משוקלל של הסתברויות הפעולה (כאשר המשקולות יכולות להתבסס על מדדי ביטחון של כל סוכן או על ביצועיו ההיסטוריים).
* **פרוטוקולי הצבעה מורכבים:** מחקר דן במיזוג מדיניויות של סוכנים בעלי תגמולים שונים באמצעות מדדי תפוסת מצב-פעולה וכללי הצבעה כמו שיטת בורדה. רעיונות אלו עשויים להיות רלוונטיים אם ארבע הסינרגיות נתפסות כתת-מדיניויות שהסוכנים "מצביעים" עליהן.
* **מיזוג סוכנים מבוסס גיוון:** גישת MARL-Focal , אף על פי שהיא מיועדת לסוכני LLM, מציעה רעיון של בחירת תת-קבוצה של סוכנים על בסיס יכולתם להשלים זה את זה, ושימוש בפתרון קונפליקטים ליצירת פלט מאוחד.
* **Value Decomposition Networks (VDN):** כפי שהוזכר קודם, VDN ושיטות דומות (למשל, QMIX) מבצעות סוג של אגרגציית פונקציות ערך כדי להעריך את הפעולה המשותפת האופטימלית.

האינטראקציה בין חלוקת נתונים אקראית, מודעות למשטר מבוססת MMD, מדיניות משותפת מורכבת (4 סינרגיות), ורעיונות סופרפוזיציה לעמידות סוכנים, יוצרת מערכת שבה התנהגות אמרגנטית (emergent behavior) היא סבירה ביותר. התנהגות זו יכולה להיות מועילה (אסטרטגיות חדשניות ומסתגלות) או מזיקה (כשלים בלתי צפויים, פעולות כאוטיות). מפת הדרכים ליישום חייבת לכלול סימולציה מקיפה ובדיקות "ארגז חול" (sandboxing) כדי להבין תכונות אמרגנטיות אלו לפני כל פריסה חיה. שלב אגרגציית המדיניות (6.3) הופך לקריטי אם יש צורך לשלב את הפלטים האישיים של הסוכנים, במיוחד אם רעיון ה"סופרפוזיציה" מוביל לפלטים הסתברותיים או למספר פעולות מועמדות מכל סוכן.

### **6.4. שיקולי אדם בלולאה (HITL)**

עבור מערכת מורכבת ואוטונומית כזו, שילוב אדם בלולאה (Human-in-the-Loop - HITL) יכול להיות בעל ערך רב. תפקידי ה-HITL יכולים לכלול:

* **אימות זיהוי משטרים:** בדיקה אנושית של זיהויי משטר שבוצעו על ידי מנוע ה-MMD, במיוחד בתקופות של אי-ודאות גבוהה.
* **פיקוח על פריסת האסטרטגיה:** מעקב אחר פעולות המערכת בזמן אמת (או בסימולציה קרובה לזמן אמת) כדי לזהות התנהגויות חריגות.
* **התערבות במצבים בלתי צפויים:** מתן אפשרות למפעיל אנושי להשבית את המערכת או לשנות את פרמטרי הפעולה שלה במקרה של תקלות או אירועי "ברבור שחור" בשוק.
* **הנחיית עידון המדיניות:** מומחה אנושי יכול לסייע בכוונון עדין של ארבע הסינרגיות או של פונקציית התגמול, על סמך ביצועי המערכת והבנתו את השוק.

ככל שהמערכת הופכת לאוטונומית יותר ומסוגלת להתנהגות אמרגנטית מורכבת (שעשויה להיות מועצמת על ידי אלמנטים בהשראת קוונטים ו-MARL), שאלות של שליטה, אחריות, ומסחר אתי הופכות למרכזיות, במיוחד בשווקים מפוקחים כמו שוק החוזים העתידיים. רכיב ה-HITL אינו רק תכונה, אלא הכרח פוטנציאלי לתפעול בטוח ואחראי. מחקר מעמיק צריך להתייחס גם למסגרת הממשל הדרושה למערכת כזו, כולל אימות מודלים, פרשנות (שהיא קשה ב-MARL), ואמצעי הגנה תפעוליים.

## **פרק 7: מסקנות והנחיות לעתיד**

המחקר הנוכחי הציג תוכנית מקיפה לפיתוח מערכת מסחר מרובת סוכנים (MARL) מתקדמת, המיועדת למסחר בחוזים עתידיים תוך שימוש בנתוני רמה 2 של רית'מיק. המערכת משלבת מספר טכנולוגיות חזיתיות, כולל מנוע זיהוי משטרי שוק מבוסס Higher-Rank MMD, ארכיטקטורת MARL שיתופית עם מדיניות מסחר מאוחדת המכילה ארבע סינרגיות, ורעיונות בהשראת סופרפוזיציה קוונטית לשיפור עמידות המערכת.

**פוטנציאל המערכת ואתגריה המרכזיים:** הפוטנציאל של מערכת כזו טמון ביכולתה לעבד מידע שוק מורכב ועשיר (נתוני רמה 2) באופן מבוזר, תוך שמירה על קוהרנטיות אסטרטגית גלובלית. זיהוי משטרי השוק נועד לאפשר למערכת להסתגל לתנאי שוק משתנים, בעוד שרעיון הסופרפוזיציה שואף להקנות לה עמידות נוספת בפני רעש ואי-ודאות. אם תמומש בהצלחה, מערכת זו עשויה להציע ביצועים עדיפים על פני גישות מסחר מסורתיות או מודלים של סוכן יחיד, במיוחד בסביבות שוק דינמיות ותנודתיות.

עם זאת, מימוש המערכת כרוך בהתמודדות עם אתגרים משמעותיים. אלו כוללים את המורכבות הטכנית של כל רכיב בנפרד (MMD מסדר גבוה, MARL, סופרפוזיציה בהשראת קוונטים), ואת הקושי שבשילובם למערכת אחת קוהרנטית ויעילה. חלוקת הנתונים האקראית למטריצות עבור הסוכנים השונים היא נקודה קריטית, הדורשת איזון עדין בין מתן "תצוגות" מגוונות לסוכנים לבין שמירה על מידע מספק בכל "תצוגה". הבטחת "שיתוף פעולה מושלם" בין סוכנים הפועלים על סמך תצפיות חלקיות היא אתגר נוסף, וכך גם עיצוב פונקציית תגמול שתאזן בין יעדי רווח והפסד, ניהול סיכונים, והימנעות מהתאמת יתר.

**היתכנות על בסיס מחקר וכלים קיימים:** למרות האתגרים, הרעיונות המוצגים במערכת נשענים על מחקרים קיימים וכלים זמינים. ספריות כמו Ray RLlib ו-PettingZoo מספקות תשתית חזקה לפיתוח ואימון מערכות MARL. מחקרים בתחום זיהוי משטרי שוק באמצעות MMD וחתימות נתיבים מציעים גישות מתקדמות הניתנות למימוש (ואף קוד פתוח זמין). הרעיון של סופרפוזיציה בהשראת קוונטים, אף שהוא חדשני, ניתן לתרגום לטכניקות RL קלאסיות כמו אנסמבלים או מדיניויות הסתברותיות, אשר נחקרו בהקשרים של עמידות והכללה. לפיכך, בעוד שהמשימה שאפתנית, היא אינה בלתי אפשרית וניתן להתקדם לקראת מימושה באופן הדרגתי ואיטרטיבי, כפי שהוצע במפת הדרכים.

**הנחיות למחקר עתידי:** הפיתוח והבדיקה של מערכת זו יניבו ככל הנראה תובנות רבות שיכולות להנחות מחקר עתידי:

1. **טכניקות MMD מתקדמות יותר:** חקירת גרסאות נוספות של MMD, אולי בשילוב עם שיטות למידה עמוקה נוספות, לזיהוי משטרים מדויק ומהיר יותר.
2. **חלוקת נתונים דינמית או למידה:** במקום חלוקה אקראית קבועה, ניתן לחקור שיטות בהן חלוקת הנתונים לסוכנים היא דינמית ומשתנה בהתאם לתנאי השוק, או אף נלמדת על ידי המערכת עצמה.
3. **אבולוציה של סינרגיות:** האפשרות שהמערכת תלמד או תפתח סינרגיות חדשות באופן אוטונומי, מעבר לארבע הסינרגיות המוגדרות מראש.
4. **שילוב עמוק יותר של עקרונות בהשראת קוונטים:** ככל שהבנתנו את האנלוגיות בין חישוב קוונטי ללמידה קלאסית תעמיק, וככל שחומרת ותוכנת החישוב הקלאסי יתפתחו, ייתכן שניתן יהיה לממש רעיונות "קוונטיים" מורכבים יותר באופן קלאסי לשיפור ביצועי סוכני RL.
5. **פרשנות ושקיפות:** פיתוח כלים וטכניקות להבנת תהליך קבלת ההחלטות של מערכת MARL מורכבת זו, במטרה להגביר את האמון בה ולהקל על איתור וטיפול בכשלים.
6. **למידה מתמשכת והסתגלות ארוכת-טווח:** חקירת ארכיטקטורות ומנגנוני למידה המאפשרים למערכת כולה להסתגל באופן רציף לשינויים ארוכי-טווח בדינמיקת השוק, מעבר לזיהוי משטרים טקטי.

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. ההתקדמות ההדרגתית, תוך התבססות על מחקר קיים וכלים מתקדמים, והתמודדות שיטתית עם האתגרים, היא המפתח להפיכת חזון זה למציאות.

Sources used in the report



[he.wikipedia.org](https://he.wikipedia.org/wiki/%D7%9C%D7%9E%D7%99%D7%93%D7%AA_%D7%97%D7%99%D7%96%D7%95%D7%A7_%D7%9E%D7%A8%D7%95%D7%91%D7%AA_%D7%A1%D7%95%D7%9B%D7%A0%D7%99%D7%9D)

[למידת חיזוק מרובת סוכנים - ויקיפדיה](https://he.wikipedia.org/wiki/%D7%9C%D7%9E%D7%99%D7%93%D7%AA_%D7%97%D7%99%D7%96%D7%95%D7%A7_%D7%9E%D7%A8%D7%95%D7%91%D7%AA_%D7%A1%D7%95%D7%9B%D7%A0%D7%99%D7%9D)

[Opens in a new window](https://he.wikipedia.org/wiki/%D7%9C%D7%9E%D7%99%D7%93%D7%AA_%D7%97%D7%99%D7%96%D7%95%D7%A7_%D7%9E%D7%A8%D7%95%D7%91%D7%AA_%D7%A1%D7%95%D7%9B%D7%A0%D7%99%D7%9D)

[](https://centerpointsecurities.com/how-to-interpret-level-2-data/)

[centerpointsecurities.com](https://centerpointsecurities.com/how-to-interpret-level-2-data/)

[How to Interpret Level 2 Data - A Complete Guide - CenterPoint Securities](https://centerpointsecurities.com/how-to-interpret-level-2-data/)

[Opens in a new window](https://centerpointsecurities.com/how-to-interpret-level-2-data/)

[](https://strategyquant.com/blog/understanding-market-regimes-indicators-in-strategyquant-coding-base/)

[strategyquant.com](https://strategyquant.com/blog/understanding-market-regimes-indicators-in-strategyquant-coding-base/)

[Understanding Market Regimes indicators in StrategyQuant Coding Base](https://strategyquant.com/blog/understanding-market-regimes-indicators-in-strategyquant-coding-base/)

[Opens in a new window](https://strategyquant.com/blog/understanding-market-regimes-indicators-in-strategyquant-coding-base/)

[](https://optimusfutures.com/blog/level-ii-market-data-on-dom/)

[optimusfutures.com](https://optimusfutures.com/blog/level-ii-market-data-on-dom/)

[How to Read Level II Market Data | Analyzing Trading Activity on the DOM - Optimus Futures](https://optimusfutures.com/blog/level-ii-market-data-on-dom/)

[Opens in a new window](https://optimusfutures.com/blog/level-ii-market-data-on-dom/)

[](https://alphaarchitect.com/regime-detection/)

[alphaarchitect.com](https://alphaarchitect.com/regime-detection/)

[A New Approach to Regime Detection and Factor Timing - Alpha Architect](https://alphaarchitect.com/regime-detection/)

[Opens in a new window](https://alphaarchitect.com/regime-detection/)

[](https://www.mdpi.com/2079-9292/14/4/820)

[mdpi.com](https://www.mdpi.com/2079-9292/14/4/820)

[A Review of Multi-Agent Reinforcement Learning Algorithms - MDPI](https://www.mdpi.com/2079-9292/14/4/820)

[Opens in a new window](https://www.mdpi.com/2079-9292/14/4/820)

[](https://www.supermarker.themarker.com/Investments/WhatAreFuturesAndWhyTradingThem.aspx)

[supermarker.themarker.com](https://www.supermarker.themarker.com/Investments/WhatAreFuturesAndWhyTradingThem.aspx)

[מהו חוזה עתידי ומה השימוש בו - Supermarker](https://www.supermarker.themarker.com/Investments/WhatAreFuturesAndWhyTradingThem.aspx)

[Opens in a new window](https://www.supermarker.themarker.com/Investments/WhatAreFuturesAndWhyTradingThem.aspx)

[](https://he.wikipedia.org/wiki/%D7%97%D7%95%D7%96%D7%94_%D7%A2%D7%AA%D7%99%D7%93%D7%99)

[he.wikipedia.org](https://he.wikipedia.org/wiki/%D7%97%D7%95%D7%96%D7%94_%D7%A2%D7%AA%D7%99%D7%93%D7%99)

[חוזה עתידי - ויקיפדיה](https://he.wikipedia.org/wiki/%D7%97%D7%95%D7%96%D7%94_%D7%A2%D7%AA%D7%99%D7%93%D7%99)

[Opens in a new window](https://he.wikipedia.org/wiki/%D7%97%D7%95%D7%96%D7%94_%D7%A2%D7%AA%D7%99%D7%93%D7%99)

[](https://www.moomoo.com/us/learn/detail-how-to-read-level-2-market-data-using-an-order-book-for-trading-strategies-66157-220709059)

[moomoo.com](https://www.moomoo.com/us/learn/detail-how-to-read-level-2-market-data-using-an-order-book-for-trading-strategies-66157-220709059)

[How to read level 2 market data - using an order book for trading strategies - Moomoo](https://www.moomoo.com/us/learn/detail-how-to-read-level-2-market-data-using-an-order-book-for-trading-strategies-66157-220709059)

[Opens in a new window](https://www.moomoo.com/us/learn/detail-how-to-read-level-2-market-data-using-an-order-book-for-trading-strategies-66157-220709059)

[](https://fiveable.me/quantum-machine-learning/unit-15/quantum-approaches-reinforcement-learning/study-guide/GKhgVq2DdbUVPU4u)

[fiveable.me](https://fiveable.me/quantum-machine-learning/unit-15/quantum-approaches-reinforcement-learning/study-guide/GKhgVq2DdbUVPU4u)

[15.2 Quantum Approaches to Reinforcement Learning - Fiveable](https://fiveable.me/quantum-machine-learning/unit-15/quantum-approaches-reinforcement-learning/study-guide/GKhgVq2DdbUVPU4u)

[Opens in a new window](https://fiveable.me/quantum-machine-learning/unit-15/quantum-approaches-reinforcement-learning/study-guide/GKhgVq2DdbUVPU4u)

[](https://arxiv.org/html/2412.18208v1)

[arxiv.org](https://arxiv.org/html/2412.18208v1)

[Quantum framework for Reinforcement Learning: integrating Markov Decision Process, quantum arithmetic, and trajectory search - arXiv](https://arxiv.org/html/2412.18208v1)

[Opens in a new window](https://arxiv.org/html/2412.18208v1)

[](https://github.com/ray-project/ray/blob/master/python/ray/tune/tune.py)

[github.com](https://github.com/ray-project/ray/blob/master/python/ray/tune/tune.py)

[ray/python/ray/tune/tune.py at master · ray-project/ray · GitHub](https://github.com/ray-project/ray/blob/master/python/ray/tune/tune.py)

[Opens in a new window](https://github.com/ray-project/ray/blob/master/python/ray/tune/tune.py)

[](https://www.investopedia.com/terms/l/level2.asp)

[investopedia.com](https://www.investopedia.com/terms/l/level2.asp)

[What Is Level 2? Definition on Nasdaq, How It Works, and Example](https://www.investopedia.com/terms/l/level2.asp)

[Opens in a new window](https://www.investopedia.com/terms/l/level2.asp)

[](https://github.com/AI4Finance-Foundation/FinRL)

[github.com](https://github.com/AI4Finance-Foundation/FinRL)

[AI4Finance-Foundation/FinRL: FinRL®: Financial ... - GitHub](https://github.com/AI4Finance-Foundation/FinRL)

[Opens in a new window](https://github.com/AI4Finance-Foundation/FinRL)

[](https://docs.copilotkit.ai/coagents/human-in-the-loop)

[docs.copilotkit.ai](https://docs.copilotkit.ai/coagents/human-in-the-loop)

[Human in the Loop (HITL) - CopilotKit](https://docs.copilotkit.ai/coagents/human-in-the-loop)

[Opens in a new window](https://docs.copilotkit.ai/coagents/human-in-the-loop)

[](https://www.anyscale.com/product/library/ray-rllib)

[anyscale.com](https://www.anyscale.com/product/library/ray-rllib)

[Ray RLlib ML Library for Reinforcement Learning - Anyscale](https://www.anyscale.com/product/library/ray-rllib)

[Opens in a new window](https://www.anyscale.com/product/library/ray-rllib)

[](https://github.com/issaz/signature-regime-detection)

[github.com](https://github.com/issaz/signature-regime-detection)

[issaz/signature-regime-detection: Code accompanying the ... - GitHub](https://github.com/issaz/signature-regime-detection)

[Opens in a new window](https://github.com/issaz/signature-regime-detection)

[](https://arxiv.org/abs/2306.15835)

[arxiv.org](https://arxiv.org/abs/2306.15835)

[arxiv.org](https://arxiv.org/abs/2306.15835)

[Opens in a new window](https://arxiv.org/abs/2306.15835)

[](https://pettingzoo.farama.org/index.html)

[pettingzoo.farama.org](https://pettingzoo.farama.org/index.html)

[PettingZoo Documentation](https://pettingzoo.farama.org/index.html)

[Opens in a new window](https://pettingzoo.farama.org/index.html)

[](https://github.com/Farama-Foundation/PettingZoo)

[github.com](https://github.com/Farama-Foundation/PettingZoo)

[Farama-Foundation/PettingZoo: An API standard for multi ... - GitHub](https://github.com/Farama-Foundation/PettingZoo)

[Opens in a new window](https://github.com/Farama-Foundation/PettingZoo)

[](https://arxiv.org/abs/2412.18208)

[arxiv.org](https://arxiv.org/abs/2412.18208)

[arxiv.org](https://arxiv.org/abs/2412.18208)

[Opens in a new window](https://arxiv.org/abs/2412.18208)

[](https://whatisai.co.il/%D7%9E%D7%94%D7%95-%D7%9E%D7%97%D7%A9%D7%95%D7%91-%D7%A7%D7%95%D7%95%D7%A0%D7%98%D7%99-%D7%91-ai/)

[whatisai.co.il](https://whatisai.co.il/%D7%9E%D7%94%D7%95-%D7%9E%D7%97%D7%A9%D7%95%D7%91-%D7%A7%D7%95%D7%95%D7%A0%D7%98%D7%99-%D7%91-ai/)

[מהו מחשוב קוונטי ב-AI? - בינה מלאכותית](https://whatisai.co.il/%D7%9E%D7%94%D7%95-%D7%9E%D7%97%D7%A9%D7%95%D7%91-%D7%A7%D7%95%D7%95%D7%A0%D7%98%D7%99-%D7%91-ai/)

[Opens in a new window](https://whatisai.co.il/%D7%9E%D7%94%D7%95-%D7%9E%D7%97%D7%A9%D7%95%D7%91-%D7%A7%D7%95%D7%95%D7%A0%D7%98%D7%99-%D7%91-ai/)

[](https://he.wikipedia.org/wiki/%D7%A1%D7%95%D7%A4%D7%A8%D7%A4%D7%95%D7%96%D7%99%D7%A6%D7%99%D7%94_%D7%A7%D7%95%D7%95%D7%A0%D7%98%D7%99%D7%AA)

[he.wikipedia.org](https://he.wikipedia.org/wiki/%D7%A1%D7%95%D7%A4%D7%A8%D7%A4%D7%95%D7%96%D7%99%D7%A6%D7%99%D7%94_%D7%A7%D7%95%D7%95%D7%A0%D7%98%D7%99%D7%AA)

[סופרפוזיציה קוונטית - ויקיפדיה](https://he.wikipedia.org/wiki/%D7%A1%D7%95%D7%A4%D7%A8%D7%A4%D7%95%D7%96%D7%99%D7%A6%D7%99%D7%94_%D7%A7%D7%95%D7%95%D7%A0%D7%98%D7%99%D7%AA)

[Opens in a new window](https://he.wikipedia.org/wiki/%D7%A1%D7%95%D7%A4%D7%A8%D7%A4%D7%95%D7%96%D7%99%D7%A6%D7%99%D7%94_%D7%A7%D7%95%D7%95%D7%A0%D7%98%D7%99%D7%AA)

[](https://machinelearningmastery.com/how-to-perform-scikit-learn-hyperparameter-optimization-with-optuna/)

[machinelearningmastery.com](https://machinelearningmastery.com/how-to-perform-scikit-learn-hyperparameter-optimization-with-optuna/)

[How to Perform Scikit-learn Hyperparameter Optimization with Optuna - MachineLearningMastery.com](https://machinelearningmastery.com/how-to-perform-scikit-learn-hyperparameter-optimization-with-optuna/)

[Opens in a new window](https://machinelearningmastery.com/how-to-perform-scikit-learn-hyperparameter-optimization-with-optuna/)

[](https://arxiv.org/html/2505.03288v1)

[arxiv.org](https://arxiv.org/html/2505.03288v1)

[Multi-Agent Deep Reinforcement Learning for Zonal Ancillary Market Coupling - arXiv](https://arxiv.org/html/2505.03288v1)

[Opens in a new window](https://arxiv.org/html/2505.03288v1)

[](https://neptune.ai/blog/how-to-optimize-hyperparameter-search)

[neptune.ai](https://neptune.ai/blog/how-to-optimize-hyperparameter-search)

[How to Optimize Hyperparameter Search Using Bayesian Optimization and Optuna](https://neptune.ai/blog/how-to-optimize-hyperparameter-search)

[Opens in a new window](https://neptune.ai/blog/how-to-optimize-hyperparameter-search)

[](https://pettingzoo.farama.org/tutorials/rllib/index.html)

[pettingzoo.farama.org](https://pettingzoo.farama.org/tutorials/rllib/index.html)

[Ray RLlib Tutorial - PettingZoo Documentation](https://pettingzoo.farama.org/tutorials/rllib/index.html)

[Opens in a new window](https://pettingzoo.farama.org/tutorials/rllib/index.html)

[](https://chuacheowhuan.github.io/MARL_CDA_env/)

[chuacheowhuan.github.io](https://chuacheowhuan.github.io/MARL_CDA_env/)

[Custom MARL (multi-agent reinforcement learning) CDA ...](https://chuacheowhuan.github.io/MARL_CDA_env/)

[Opens in a new window](https://chuacheowhuan.github.io/MARL_CDA_env/)

[](https://3commas.io/blog/comprehensive-2025-guide-to-backtesting-ai-trading)

[3commas.io](https://3commas.io/blog/comprehensive-2025-guide-to-backtesting-ai-trading)

[Comprehensive 2025 Guide to Backtesting AI Crypto Trading Strategies - 3Commas](https://3commas.io/blog/comprehensive-2025-guide-to-backtesting-ai-trading)

[Opens in a new window](https://3commas.io/blog/comprehensive-2025-guide-to-backtesting-ai-trading)

[](http://www.arxiv.org/pdf/2408.10932)

[arxiv.org](http://www.arxiv.org/pdf/2408.10932)

[The Evolution of Reinforcement Learning in Quantitative Finance: A Survey - arXiv](http://www.arxiv.org/pdf/2408.10932)

[Opens in a new window](http://www.arxiv.org/pdf/2408.10932)

[](https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0236178)

[journals.plos.org](https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0236178)

[Action-specialized expert ensemble trading system with extended discrete action space using deep reinforcement learning | PLOS One](https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0236178)

[Opens in a new window](https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0236178)

[](https://arxiv.org/html/2404.06387v2)

[arxiv.org](https://arxiv.org/html/2404.06387v2)

[Robust Coordination under Misaligned Communication via Power Regularization - arXiv](https://arxiv.org/html/2404.06387v2)

[Opens in a new window](https://arxiv.org/html/2404.06387v2)

[](https://arxiv.org/html/2412.15388v1)

[arxiv.org](https://arxiv.org/html/2412.15388v1)

[Investigating Relational State Abstraction in Collaborative MARL - arXiv](https://arxiv.org/html/2412.15388v1)

[Opens in a new window](https://arxiv.org/html/2412.15388v1)

[](https://seqml.github.io/marl4fin/)

[seqml.github.io](https://seqml.github.io/marl4fin/)

[Learning Multi-Agent Intention-Aware Communication for Optimal Multi-Order Execution in Finance](https://seqml.github.io/marl4fin/)

[Opens in a new window](https://seqml.github.io/marl4fin/)

[](https://su.diva-portal.org/smash/get/diva2:1955587/FULLTEXT01.pdf)

[su.diva-portal.org](https://su.diva-portal.org/smash/get/diva2:1955587/FULLTEXT01.pdf)

[Enhancing Multi-Agent Deep Reinforcement Learning (MADRL) for Financial Trading - DiVA portal](https://su.diva-portal.org/smash/get/diva2:1955587/FULLTEXT01.pdf)

[Opens in a new window](https://su.diva-portal.org/smash/get/diva2:1955587/FULLTEXT01.pdf)

[](https://alphaarchitect.com/reinforcement-learning-for-trading/)

[alphaarchitect.com](https://alphaarchitect.com/reinforcement-learning-for-trading/)

[Deep Reinforcement Learning for trading applications - Alpha Architect](https://alphaarchitect.com/reinforcement-learning-for-trading/)

[Opens in a new window](https://alphaarchitect.com/reinforcement-learning-for-trading/)

[](https://github.com/optuna/optuna-examples/blob/main/ray/ray_joblib.py)

[github.com](https://github.com/optuna/optuna-examples/blob/main/ray/ray_joblib.py)

[optuna-examples/ray/ray\_joblib.py at main - GitHub](https://github.com/optuna/optuna-examples/blob/main/ray/ray_joblib.py)

[Opens in a new window](https://github.com/optuna/optuna-examples/blob/main/ray/ray_joblib.py)

[](https://github.com/LSEG-API-Samples/Article.RD.Python.MarketRegimeDetectionUsingStatisticalAndMLBasedApproaches/blob/main/Market%20regime%20detection.ipynb)

[github.com](https://github.com/LSEG-API-Samples/Article.RD.Python.MarketRegimeDetectionUsingStatisticalAndMLBasedApproaches/blob/main/Market%20regime%20detection.ipynb)

[Market regime detection.ipynb - GitHub](https://github.com/LSEG-API-Samples/Article.RD.Python.MarketRegimeDetectionUsingStatisticalAndMLBasedApproaches/blob/main/Market%20regime%20detection.ipynb)

[Opens in a new window](https://github.com/LSEG-API-Samples/Article.RD.Python.MarketRegimeDetectionUsingStatisticalAndMLBasedApproaches/blob/main/Market%20regime%20detection.ipynb)

[](https://arxiv.org/html/2407.10403v1)

[arxiv.org](https://arxiv.org/html/2407.10403v1)

[Cooperative Reward Shaping for Multi-Agent Pathfinding - arXiv](https://arxiv.org/html/2407.10403v1)

[Opens in a new window](https://arxiv.org/html/2407.10403v1)

[](https://docs.ray.io/en/latest/tune/examples/optuna_example.html)

[docs.ray.io](https://docs.ray.io/en/latest/tune/examples/optuna_example.html)

[Running Tune experiments with Optuna - Ray Docs](https://docs.ray.io/en/latest/tune/examples/optuna_example.html)

[Opens in a new window](https://docs.ray.io/en/latest/tune/examples/optuna_example.html)

[](https://rl.airlab.deib.polimi.it/wp-content/uploads/2022/10/ICAIF_2022___Expert_Learning-1.pdf)

[rl.airlab.deib.polimi.it](https://rl.airlab.deib.polimi.it/wp-content/uploads/2022/10/ICAIF_2022___Expert_Learning-1.pdf)

[Addressing Non-Stationarity in FX Trading with Online Model Selection of Offline RL Experts - RL@Polimi - Politecnico di Milano](https://rl.airlab.deib.polimi.it/wp-content/uploads/2022/10/ICAIF_2022___Expert_Learning-1.pdf)

[Opens in a new window](https://rl.airlab.deib.polimi.it/wp-content/uploads/2022/10/ICAIF_2022___Expert_Learning-1.pdf)

[](https://github.com/yvesdhondt/MarketMoodRing)

[github.com](https://github.com/yvesdhondt/MarketMoodRing)

[yvesdhondt/MarketMoodRing: Python implementation for regime-dependent portfolio optimization - GitHub](https://github.com/yvesdhondt/MarketMoodRing)

[Opens in a new window](https://github.com/yvesdhondt/MarketMoodRing)

[](https://www.ri.cmu.edu/app/uploads/2024/12/Thesis-Yeeho-Song-Final.pdf)

[ri.cmu.edu](https://www.ri.cmu.edu/app/uploads/2024/12/Thesis-Yeeho-Song-Final.pdf)

[Robust Reinforcement Learning via Curricular Learning - Carnegie Mellon University's Robotics Institute](https://www.ri.cmu.edu/app/uploads/2024/12/Thesis-Yeeho-Song-Final.pdf)

[Opens in a new window](https://www.ri.cmu.edu/app/uploads/2024/12/Thesis-Yeeho-Song-Final.pdf)

[](https://www.researchgate.net/publication/23249669_Quantum_Reinforcement_Learning)

[researchgate.net](https://www.researchgate.net/publication/23249669_Quantum_Reinforcement_Learning)

[(PDF) Quantum Reinforcement Learning - ResearchGate](https://www.researchgate.net/publication/23249669_Quantum_Reinforcement_Learning)

[Opens in a new window](https://www.researchgate.net/publication/23249669_Quantum_Reinforcement_Learning)

[](https://philarchive.org/archive/TALQMM)

[philarchive.org](https://philarchive.org/archive/TALQMM)

[Merging quantum computing with next-gen AI - PhilArchive](https://philarchive.org/archive/TALQMM)

[Opens in a new window](https://philarchive.org/archive/TALQMM)

[](https://arxiv.org/abs/2502.17518)

[arxiv.org](https://arxiv.org/abs/2502.17518)

[[2502.17518] Ensemble RL through Classifier Models: Enhancing Risk-Return Trade-offs in Trading Strategies - arXiv](https://arxiv.org/abs/2502.17518)

[Opens in a new window](https://arxiv.org/abs/2502.17518)

[](https://www.researchgate.net/publication/390768005_Quantum-Inspired_Defenses_for_Adversarial_Robustness_Leveraging_Quantum_Computing_Concepts_to_Secure_ML_Systems)

[researchgate.net](https://www.researchgate.net/publication/390768005_Quantum-Inspired_Defenses_for_Adversarial_Robustness_Leveraging_Quantum_Computing_Concepts_to_Secure_ML_Systems)

[(PDF) Quantum-Inspired Defenses for Adversarial Robustness: Leveraging Quantum Computing Concepts to Secure ML Systems - ResearchGate](https://www.researchgate.net/publication/390768005_Quantum-Inspired_Defenses_for_Adversarial_Robustness_Leveraging_Quantum_Computing_Concepts_to_Secure_ML_Systems)

[Opens in a new window](https://www.researchgate.net/publication/390768005_Quantum-Inspired_Defenses_for_Adversarial_Robustness_Leveraging_Quantum_Computing_Concepts_to_Secure_ML_Systems)

[](https://www.researchgate.net/publication/383453522_Ensemble_Deep_Reinforcement_Learning_for_Financial_Trading)

[researchgate.net](https://www.researchgate.net/publication/383453522_Ensemble_Deep_Reinforcement_Learning_for_Financial_Trading)

[Ensemble Deep Reinforcement Learning for Financial Trading | Request PDF](https://www.researchgate.net/publication/383453522_Ensemble_Deep_Reinforcement_Learning_for_Financial_Trading)

[Opens in a new window](https://www.researchgate.net/publication/383453522_Ensemble_Deep_Reinforcement_Learning_for_Financial_Trading)

[](https://smythos.com/ai-agents/multi-agent-systems/multi-agent-systems-and-reinforcement-learning/)

[smythos.com](https://smythos.com/ai-agents/multi-agent-systems/multi-agent-systems-and-reinforcement-learning/)

[Integrating Multi-Agent Systems and Reinforcement Learning - SmythOS](https://smythos.com/ai-agents/multi-agent-systems/multi-agent-systems-and-reinforcement-learning/)

[Opens in a new window](https://smythos.com/ai-agents/multi-agent-systems/multi-agent-systems-and-reinforcement-learning/)

[](https://www.researchgate.net/publication/386279469_Multi-Agent_Reinforcement_Learning_for_High-Frequency_Trading_Strategy_Optimization)

[researchgate.net](https://www.researchgate.net/publication/386279469_Multi-Agent_Reinforcement_Learning_for_High-Frequency_Trading_Strategy_Optimization)

[Multi-Agent Reinforcement Learning for High-Frequency Trading Strategy Optimization](https://www.researchgate.net/publication/386279469_Multi-Agent_Reinforcement_Learning_for_High-Frequency_Trading_Strategy_Optimization)

[Opens in a new window](https://www.researchgate.net/publication/386279469_Multi-Agent_Reinforcement_Learning_for_High-Frequency_Trading_Strategy_Optimization)

[](https://docs.ray.io/en/latest/rllib/multi-agent-envs.html)

[docs.ray.io](https://docs.ray.io/en/latest/rllib/multi-agent-envs.html)

[Multi-Agent Environments - RLlib - Ray Docs](https://docs.ray.io/en/latest/rllib/multi-agent-envs.html)

[Opens in a new window](https://docs.ray.io/en/latest/rllib/multi-agent-envs.html)

[](https://weareadaptive.com/trading-resources/blog/machine-learning-models-in-finance/)

[weareadaptive.com](https://weareadaptive.com/trading-resources/blog/machine-learning-models-in-finance/)

[Machine learning models in finance | WeAreAdaptive.com](https://weareadaptive.com/trading-resources/blog/machine-learning-models-in-finance/)

[Opens in a new window](https://weareadaptive.com/trading-resources/blog/machine-learning-models-in-finance/)

[](https://papers.nips.cc/paper_files/paper/2024/file/5aee125f052c90e326dcf6f380df94f6-Paper-Datasets_and_Benchmarks_Track.pdf)

[papers.nips.cc](https://papers.nips.cc/paper_files/paper/2024/file/5aee125f052c90e326dcf6f380df94f6-Paper-Datasets_and_Benchmarks_Track.pdf)

[JaxMARL: Multi-Agent RL Environments and Algorithms in JAX - NIPS papers](https://papers.nips.cc/paper_files/paper/2024/file/5aee125f052c90e326dcf6f380df94f6-Paper-Datasets_and_Benchmarks_Track.pdf)

[Opens in a new window](https://papers.nips.cc/paper_files/paper/2024/file/5aee125f052c90e326dcf6f380df94f6-Paper-Datasets_and_Benchmarks_Track.pdf)

[](https://www.multicharts.com/trading-software/index.php?title=Walk_Forward_Optimization)

[multicharts.com](https://www.multicharts.com/trading-software/index.php?title=Walk_Forward_Optimization)

[Walk Forward Optimization: Analysis & Testing Trading Systems - Help MultiCharts](https://www.multicharts.com/trading-software/index.php?title=Walk_Forward_Optimization)

[Opens in a new window](https://www.multicharts.com/trading-software/index.php?title=Walk_Forward_Optimization)

[](https://digital.kenyon.edu/cgi/viewcontent.cgi?article=1008&context=dh_iphs_ai)

[digital.kenyon.edu](https://digital.kenyon.edu/cgi/viewcontent.cgi?article=1008&context=dh_iphs_ai)

[Deep Reinforcement Learning in Trading Algorithms - Digital Kenyon](https://digital.kenyon.edu/cgi/viewcontent.cgi?article=1008&context=dh_iphs_ai)

[Opens in a new window](https://digital.kenyon.edu/cgi/viewcontent.cgi?article=1008&context=dh_iphs_ai)

[](https://algotrading101.com/learn/walk-forward-optimization/)

[algotrading101.com](https://algotrading101.com/learn/walk-forward-optimization/)

[What is a Walk-Forward Optimization and How to Run It? - Algo Trading 101](https://algotrading101.com/learn/walk-forward-optimization/)

[Opens in a new window](https://algotrading101.com/learn/walk-forward-optimization/)

[](https://www.numberanalytics.com/blog/mastering-advanced-backtesting-trading-success)

[numberanalytics.com](https://www.numberanalytics.com/blog/mastering-advanced-backtesting-trading-success)

[Mastering Advanced Backtesting for Trading Success - Number Analytics](https://www.numberanalytics.com/blog/mastering-advanced-backtesting-trading-success)

[Opens in a new window](https://www.numberanalytics.com/blog/mastering-advanced-backtesting-trading-success)

[](https://www.newtrading.io/backtesting-guide/)

[newtrading.io](https://www.newtrading.io/backtesting-guide/)

[How To Backtest A Trading Strategy: A Step-by-Step Guide - NewTrading](https://www.newtrading.io/backtesting-guide/)

[Opens in a new window](https://www.newtrading.io/backtesting-guide/)

[](https://openreview.net/forum?id=ybiUVIxJth)

[openreview.net](https://openreview.net/forum?id=ybiUVIxJth)

[Policy Aggregation - OpenReview](https://openreview.net/forum?id=ybiUVIxJth)

[Opens in a new window](https://openreview.net/forum?id=ybiUVIxJth)

[](https://arxiv.org/html/2505.11862v1)

[arxiv.org](https://arxiv.org/html/2505.11862v1)

[Q-Policy: Quantum-Enhanced Policy Evaluation for Scalable Reinforcement Learning - arXiv](https://arxiv.org/html/2505.11862v1)

[Opens in a new window](https://arxiv.org/html/2505.11862v1)

[](https://www.quandela.com/resources/blog/quantum-reinforcement-learning-gets-an-entanglement-boost-using-single-photons/)

[quandela.com](https://www.quandela.com/resources/blog/quantum-reinforcement-learning-gets-an-entanglement-boost-using-single-photons/)

[Quantum Reinforcement Learning Gets an Entanglement Boost Using Single Photons - Quandela](https://www.quandela.com/resources/blog/quantum-reinforcement-learning-gets-an-entanglement-boost-using-single-photons/)

[Opens in a new window](https://www.quandela.com/resources/blog/quantum-reinforcement-learning-gets-an-entanglement-boost-using-single-photons/)

[](https://arxiv.org/html/2303.02618v3)

[arxiv.org](https://arxiv.org/html/2303.02618v3)

[Ensemble Reinforcement Learning: A Survey - arXiv](https://arxiv.org/html/2303.02618v3)

[Opens in a new window](https://arxiv.org/html/2303.02618v3)

[](https://www.researchgate.net/publication/221093843_Random_Feature_Subset_Selection_for_Ensemble_Based_Classification_of_Data_with_Missing_Features)

[researchgate.net](https://www.researchgate.net/publication/221093843_Random_Feature_Subset_Selection_for_Ensemble_Based_Classification_of_Data_with_Missing_Features)

[Random Feature Subset Selection for Ensemble Based Classification of Data with Missing Features - ResearchGate](https://www.researchgate.net/publication/221093843_Random_Feature_Subset_Selection_for_Ensemble_Based_Classification_of_Data_with_Missing_Features)

[Opens in a new window](https://www.researchgate.net/publication/221093843_Random_Feature_Subset_Selection_for_Ensemble_Based_Classification_of_Data_with_Missing_Features)

[](https://arxiv.org/pdf/2306.15835)

[arxiv.org](https://arxiv.org/pdf/2306.15835)

[arXiv:2306.15835v1 [stat.ML] 27 Jun 2023](https://arxiv.org/pdf/2306.15835)

[Opens in a new window](https://arxiv.org/pdf/2306.15835)

[](https://openreview.net/pdf?id=oAkSRhl3qU)

[openreview.net](https://openreview.net/pdf?id=oAkSRhl3qU)

[Mitigating Relative Over-Generalization in Multi-Agent Re- inforcement Learning - OpenReview](https://openreview.net/pdf?id=oAkSRhl3qU)

[Opens in a new window](https://openreview.net/pdf?id=oAkSRhl3qU)

[](https://openreview.net/forum?id=Cfbr56K4gp)

[openreview.net](https://openreview.net/forum?id=Cfbr56K4gp)

[Robust RLHF with Noisy Rewards - OpenReview](https://openreview.net/forum?id=Cfbr56K4gp)

[Opens in a new window](https://openreview.net/forum?id=Cfbr56K4gp)

[](https://arxiv.org/pdf/2411.08637?)

[arxiv.org](https://arxiv.org/pdf/2411.08637?)

[Robot See, Robot Do: Imitation Reward for Noisy Financial Environments - arXiv](https://arxiv.org/pdf/2411.08637?)

[Opens in a new window](https://arxiv.org/pdf/2411.08637?)

[](https://www.lamda.nju.edu.cn/lilh/file/semdiv.pdf)

[lamda.nju.edu.cn](https://www.lamda.nju.edu.cn/lilh/file/semdiv.pdf)

[LLM-Assisted Semantically Diverse Teammates Generationfor Efficient Multi-agent Coordination - LAMDA](https://www.lamda.nju.edu.cn/lilh/file/semdiv.pdf)

[Opens in a new window](https://www.lamda.nju.edu.cn/lilh/file/semdiv.pdf)

[](https://japmi.org/index.php/japmi/article/download/13/13/28)

[japmi.org](https://japmi.org/index.php/japmi/article/download/13/13/28)

[Multi-Agent Reinforcement Learning for High-Frequency Trading Strategy Optimization](https://japmi.org/index.php/japmi/article/download/13/13/28)

[Opens in a new window](https://japmi.org/index.php/japmi/article/download/13/13/28)

[](https://arxiv.org/abs/2505.08630)

[arxiv.org](https://arxiv.org/abs/2505.08630)

[[2505.08630] Credit Assignment and Efficient Exploration based on Influence Scope in Multi-agent Reinforcement Learning - arXiv](https://arxiv.org/abs/2505.08630)

[Opens in a new window](https://arxiv.org/abs/2505.08630)

[](https://openreview.net/forum?id=ZOTpFhwQQ9)

[openreview.net](https://openreview.net/forum?id=ZOTpFhwQQ9)

[Efficient Credit Assignment in Cooperative Multi-Agent Reinforcement Learning | OpenReview](https://openreview.net/forum?id=ZOTpFhwQQ9)

[Opens in a new window](https://openreview.net/forum?id=ZOTpFhwQQ9)

[](https://www.aalpha.net/blog/how-to-build-multi-agent-ai-system/)

[aalpha.net](https://www.aalpha.net/blog/how-to-build-multi-agent-ai-system/)

[How to Build a Multi-Agent AI System : In-Depth Guide](https://www.aalpha.net/blog/how-to-build-multi-agent-ai-system/)

[Opens in a new window](https://www.aalpha.net/blog/how-to-build-multi-agent-ai-system/)

[](https://iclr.cc/virtual/2025/events/spotlight-posters)

[iclr.cc](https://iclr.cc/virtual/2025/events/spotlight-posters)

[ICLR 2025 Spotlights](https://iclr.cc/virtual/2025/events/spotlight-posters)

[Opens in a new window](https://iclr.cc/virtual/2025/events/spotlight-posters)

[](https://www.mdpi.com/2079-9292/14/9/1721)

[mdpi.com](https://www.mdpi.com/2079-9292/14/9/1721)

[Artificial Intelligence vs. Efficient Markets: A Critical Reassessment of Predictive Models in the Big Data Era - MDPI](https://www.mdpi.com/2079-9292/14/9/1721)

[Opens in a new window](https://www.mdpi.com/2079-9292/14/9/1721)

[](https://arxiv.org/html/2412.04031v1)

[arxiv.org](https://arxiv.org/html/2412.04031v1)

[Dimension Reduction via Random Projection for Privacy in Multi-Agent Systems - arXiv](https://arxiv.org/html/2412.04031v1)

[Opens in a new window](https://arxiv.org/html/2412.04031v1)

[](https://www.researchgate.net/publication/344436840_PettingZoo_Gym_for_Multi-Agent_Reinforcement_Learning)

[researchgate.net](https://www.researchgate.net/publication/344436840_PettingZoo_Gym_for_Multi-Agent_Reinforcement_Learning)

[PettingZoo: Gym for Multi-Agent Reinforcement Learning - ResearchGate](https://www.researchgate.net/publication/344436840_PettingZoo_Gym_for_Multi-Agent_Reinforcement_Learning)

[Opens in a new window](https://www.researchgate.net/publication/344436840_PettingZoo_Gym_for_Multi-Agent_Reinforcement_Learning)

[](https://github.com/AI4Finance-Foundation/FinRL-Meta)

[github.com](https://github.com/AI4Finance-Foundation/FinRL-Meta)

[FinRL®-Meta: Dynamic datasets and market environments for FinRL. - GitHub](https://github.com/AI4Finance-Foundation/FinRL-Meta)

[Opens in a new window](https://github.com/AI4Finance-Foundation/FinRL-Meta)

[](https://www.religareonline.com/blog/backtesting-your-technical-trading-strategies/)

[religareonline.com](https://www.religareonline.com/blog/backtesting-your-technical-trading-strategies/)

[Backtesting Your Technical Trading Strategies - Religare Broking](https://www.religareonline.com/blog/backtesting-your-technical-trading-strategies/)

[Opens in a new window](https://www.religareonline.com/blog/backtesting-your-technical-trading-strategies/)

[](https://arxiv.org/html/2403.03185v4)

[arxiv.org](https://arxiv.org/html/2403.03185v4)

[Correlated Proxies: A New Definition and Improved Mitigation for Reward Hacking - arXiv](https://arxiv.org/html/2403.03185v4)

[Opens in a new window](https://arxiv.org/html/2403.03185v4)

[](https://arxiv.org/abs/2502.18770v1?utm_source=tldrai)

[arxiv.org](https://arxiv.org/abs/2502.18770v1?utm_source=tldrai)

[[2502.18770v1] Reward Shaping to Mitigate Reward Hacking in RLHF - arXiv](https://arxiv.org/abs/2502.18770v1?utm_source=tldrai)

[Opens in a new window](https://arxiv.org/abs/2502.18770v1?utm_source=tldrai)

[](https://www.researchgate.net/publication/391476570_Heterogeneous_Multi-Agent_Reinforcement_Learning_based_on_Modularized_Policy_Network)

[researchgate.net](https://www.researchgate.net/publication/391476570_Heterogeneous_Multi-Agent_Reinforcement_Learning_based_on_Modularized_Policy_Network)

[Heterogeneous Multi-Agent Reinforcement Learning based on Modularized Policy Network | Request PDF - ResearchGate](https://www.researchgate.net/publication/391476570_Heterogeneous_Multi-Agent_Reinforcement_Learning_based_on_Modularized_Policy_Network)

[Opens in a new window](https://www.researchgate.net/publication/391476570_Heterogeneous_Multi-Agent_Reinforcement_Learning_based_on_Modularized_Policy_Network)

[](https://proceedings.neurips.cc/paper_files/paper/2024/file/274d0146144643ee2459a602123c60ff-Paper-Conference.pdf)

[proceedings.neurips.cc](https://proceedings.neurips.cc/paper_files/paper/2024/file/274d0146144643ee2459a602123c60ff-Paper-Conference.pdf)

[Kaleidoscope: Learnable Masks for Heterogeneous Multi-agent Reinforcement Learning - NIPS papers](https://proceedings.neurips.cc/paper_files/paper/2024/file/274d0146144643ee2459a602123c60ff-Paper-Conference.pdf)

[Opens in a new window](https://proceedings.neurips.cc/paper_files/paper/2024/file/274d0146144643ee2459a602123c60ff-Paper-Conference.pdf)

[](https://arxiv.org/html/2502.03723v2)

[arxiv.org](https://arxiv.org/html/2502.03723v2)

[Speaking the Language of Teamwork: LLM-Guided Credit Assignment in Multi-Agent Reinforcement Learning - arXiv](https://arxiv.org/html/2502.03723v2)

[Opens in a new window](https://arxiv.org/html/2502.03723v2)

[](https://arxiv.org/abs/2502.04492)

[arxiv.org](https://arxiv.org/abs/2502.04492)

[[2502.04492] Multi-Agent Reinforcement Learning with Focal Diversity Optimization - arXiv](https://arxiv.org/abs/2502.04492)

[Opens in a new window](https://arxiv.org/abs/2502.04492)

[](https://arxiv.org/html/2501.10709v1)

[arxiv.org](https://arxiv.org/html/2501.10709v1)

[Revisiting Ensemble Methods for Stock Trading and Crypto Trading Tasks at ACM ICAIF FinRL Contests 2023/2024 - arXiv](https://arxiv.org/html/2501.10709v1)

[Opens in a new window](https://arxiv.org/html/2501.10709v1)

[](https://pmc.ncbi.nlm.nih.gov/articles/PMC11888913/)

[pmc.ncbi.nlm.nih.gov](https://pmc.ncbi.nlm.nih.gov/articles/PMC11888913/)

[Stock market trading via actor-critic reinforcement learning and adaptable data structure - PMC - PubMed Central](https://pmc.ncbi.nlm.nih.gov/articles/PMC11888913/)

[Opens in a new window](https://pmc.ncbi.nlm.nih.gov/articles/PMC11888913/)

[](https://intranet.csc.liv.ac.uk/~rahul/papers/Spooner_thesis.pdf)

[intranet.csc.liv.ac.uk](https://intranet.csc.liv.ac.uk/~rahul/papers/Spooner_thesis.pdf)

[Algorithmic Trading and Reinforcement Learning - University of Liverpool](https://intranet.csc.liv.ac.uk/~rahul/papers/Spooner_thesis.pdf)

[Opens in a new window](https://intranet.csc.liv.ac.uk/~rahul/papers/Spooner_thesis.pdf)

[](https://rongpeng.info/images/pdfs/2025_Geng_Noise%20distribution%20decomposition%20based%20multi-agent%20distributional%20reinforcement.pdf)

[rongpeng.info](https://rongpeng.info/images/pdfs/2025_Geng_Noise%20distribution%20decomposition%20based%20multi-agent%20distributional%20reinforcement.pdf)

[Noise Distribution Decomposition Based Multi-Agent Distributional Reinforcement Learning - Dr. Rongpeng Li](https://rongpeng.info/images/pdfs/2025_Geng_Noise%20distribution%20decomposition%20based%20multi-agent%20distributional%20reinforcement.pdf)

[Opens in a new window](https://rongpeng.info/images/pdfs/2025_Geng_Noise%20distribution%20decomposition%20based%20multi-agent%20distributional%20reinforcement.pdf)

[](https://www.ijraset.com/best-journal/quantum-inspired-dynamic-decision-making-algorithm-qiddm-a-robust-framework-for-delayed-commitment-in-uncertain-environments)

[ijraset.com](https://www.ijraset.com/best-journal/quantum-inspired-dynamic-decision-making-algorithm-qiddm-a-robust-framework-for-delayed-commitment-in-uncertain-environments)

[Quantum-Inspired Dynamic Decision-Making Algorithm (QIDDM) - IJRASET](https://www.ijraset.com/best-journal/quantum-inspired-dynamic-decision-making-algorithm-qiddm-a-robust-framework-for-delayed-commitment-in-uncertain-environments)

[Opens in a new window](https://www.ijraset.com/best-journal/quantum-inspired-dynamic-decision-making-algorithm-qiddm-a-robust-framework-for-delayed-commitment-in-uncertain-environments)

[](https://arxiv.org/pdf/2006.05574)

[arxiv.org](https://arxiv.org/pdf/2006.05574)

[Multi-Agent Reinforcement Learning in a Realistic Limit Order Book Market Simulation - arXiv](https://arxiv.org/pdf/2006.05574)

[Opens in a new window](https://arxiv.org/pdf/2006.05574)

[](https://howtotrade.com/blog/level-2-market-data/)

[howtotrade.com](https://howtotrade.com/blog/level-2-market-data/)

[How to Read Level 2 Market Data Like a Pro - HowToTrade](https://howtotrade.com/blog/level-2-market-data/)

[Opens in a new window](https://howtotrade.com/blog/level-2-market-data/)

[](https://www.quantifiedstrategies.com/backtesting/)

[quantifiedstrategies.com](https://www.quantifiedstrategies.com/backtesting/)

[Backtesting: Definition, Example, How It Works, and Downsides - QuantifiedStrategies.com](https://www.quantifiedstrategies.com/backtesting/)

[Opens in a new window](https://www.quantifiedstrategies.com/backtesting/)

[](https://forextester.com/blog/futures-backtesting)

[forextester.com](https://forextester.com/blog/futures-backtesting)

[Backtesting Futures Trading Strategies in 2025 - Forex Tester](https://forextester.com/blog/futures-backtesting)

[Opens in a new window](https://forextester.com/blog/futures-backtesting)

[](https://pettingzoo.farama.org/content/environment_creation/)

[pettingzoo.farama.org](https://pettingzoo.farama.org/content/environment_creation/)

[Environment Creation - PettingZoo Documentation](https://pettingzoo.farama.org/content/environment_creation/)

[Opens in a new window](https://pettingzoo.farama.org/content/environment_creation/)

[](https://pmc.ncbi.nlm.nih.gov/articles/PMC9823489/)

[pmc.ncbi.nlm.nih.gov](https://pmc.ncbi.nlm.nih.gov/articles/PMC9823489/)

[Feature Selection Method Using Multi-Agent Reinforcement Learning Based on Guide Agents - PMC](https://pmc.ncbi.nlm.nih.gov/articles/PMC9823489/)

[Opens in a new window](https://pmc.ncbi.nlm.nih.gov/articles/PMC9823489/)

[](https://jmlr.org/papers/volume20/18-476/18-476.pdf)

[jmlr.org](https://jmlr.org/papers/volume20/18-476/18-476.pdf)

[Deep Reinforcement Learning for Swarm Systems](https://jmlr.org/papers/volume20/18-476/18-476.pdf)

[Opens in a new window](https://jmlr.org/papers/volume20/18-476/18-476.pdf)

[](https://www.ijcai.org/proceedings/2020/0623.pdf)

[ijcai.org](https://www.ijcai.org/proceedings/2020/0623.pdf)

[MAPS: Multi-Agent reinforcement learning-based Portfolio management System. - IJCAI](https://www.ijcai.org/proceedings/2020/0623.pdf)

[Opens in a new window](https://www.ijcai.org/proceedings/2020/0623.pdf)

Sources read but not used in the report



[whatisai.co.il](https://whatisai.co.il/%D7%9E%D7%94-%D7%96%D7%94-%D7%9C%D7%9E%D7%99%D7%93%D7%AA-%D7%97%D7%99%D7%96%D7%95%D7%A7/)

[מה זה למידת חיזוק? - בינה מלאכותית](https://whatisai.co.il/%D7%9E%D7%94-%D7%96%D7%94-%D7%9C%D7%9E%D7%99%D7%93%D7%AA-%D7%97%D7%99%D7%96%D7%95%D7%A7/)

[Opens in a new window](https://whatisai.co.il/%D7%9E%D7%94-%D7%96%D7%94-%D7%9C%D7%9E%D7%99%D7%93%D7%AA-%D7%97%D7%99%D7%96%D7%95%D7%A7/)

[](https://arxiv.org/html/2402.10803v1)

[arxiv.org](https://arxiv.org/html/2402.10803v1)

[Modelling crypto markets by multi-agent reinforcement learning - arXiv](https://arxiv.org/html/2402.10803v1)

[Opens in a new window](https://arxiv.org/html/2402.10803v1)

[](https://www.ig.com/uk/trading-strategies/how-to-trade-using-level-2-shares-data-211118)

[ig.com](https://www.ig.com/uk/trading-strategies/how-to-trade-using-level-2-shares-data-211118)

[What is Level 2 Data and How Can You Trade Using it? - IG UK](https://www.ig.com/uk/trading-strategies/how-to-trade-using-level-2-shares-data-211118)

[Opens in a new window](https://www.ig.com/uk/trading-strategies/how-to-trade-using-level-2-shares-data-211118)

[](https://arxiv.org/abs/1809.04999)

[arxiv.org](https://arxiv.org/abs/1809.04999)

[arxiv.org](https://arxiv.org/abs/1809.04999)

[Opens in a new window](https://arxiv.org/abs/1809.04999)

[](https://arxiv.org/abs/0810.3828)

[arxiv.org](https://arxiv.org/abs/0810.3828)

[arxiv.org](https://arxiv.org/abs/0810.3828)

[Opens in a new window](https://arxiv.org/abs/0810.3828)

[](https://github.com/minqi/learning-to-communicate-pytorch)

[github.com](https://github.com/minqi/learning-to-communicate-pytorch)

[minqi/learning-to-communicate-pytorch: Learning to ... - GitHub](https://github.com/minqi/learning-to-communicate-pytorch)

[Opens in a new window](https://github.com/minqi/learning-to-communicate-pytorch)

[](https://arxiv.org/pdf/2307.14745)

[arxiv.org](https://arxiv.org/pdf/2307.14745)

[arxiv.org](https://arxiv.org/pdf/2307.14745)

[Opens in a new window](https://arxiv.org/pdf/2307.14745)

[](https://proceedings.mlr.press/v233/bauman24a/bauman24a.pdf)

[proceedings.mlr.press](https://proceedings.mlr.press/v233/bauman24a/bauman24a.pdf)

[proceedings.mlr.press](https://proceedings.mlr.press/v233/bauman24a/bauman24a.pdf)

[Opens in a new window](https://proceedings.mlr.press/v233/bauman24a/bauman24a.pdf)

[](https://arxiv.org/abs/2009.14471)

[arxiv.org](https://arxiv.org/abs/2009.14471)

[arxiv.org](https://arxiv.org/abs/2009.14471)

[Opens in a new window](https://arxiv.org/abs/2009.14471)

[](https://arxiv.org/abs/1812.11794)

[arxiv.org](https://arxiv.org/abs/1812.11794)

[arxiv.org](https://arxiv.org/abs/1812.11794)

[Opens in a new window](https://arxiv.org/abs/1812.11794)

[](https://arxiv.org/abs/2405.11778)

[arxiv.org](https://arxiv.org/abs/2405.11778)

[arxiv.org](https://arxiv.org/abs/2405.11778)

[Opens in a new window](https://arxiv.org/abs/2405.11778)

[](https://arxiv.org/abs/2312.10256)

[arxiv.org](https://arxiv.org/abs/2312.10256)

[arxiv.org](https://arxiv.org/abs/2312.10256)

[Opens in a new window](https://arxiv.org/abs/2312.10256)

[](https://optuna.org/)

[optuna.org](https://optuna.org/)

[Optuna - A hyperparameter optimization framework](https://optuna.org/)

[Opens in a new window](https://optuna.org/)

[](https://academic.oup.com/jrsssb/article/79/4/959/7040648)

[academic.oup.com](https://academic.oup.com/jrsssb/article/79/4/959/7040648)

[Random-projection Ensemble Classification | Journal of the Royal Statistical Society Series B - Oxford Academic](https://academic.oup.com/jrsssb/article/79/4/959/7040648)

[Opens in a new window](https://academic.oup.com/jrsssb/article/79/4/959/7040648)

[](https://www.researchgate.net/publication/377136235_Random_projection_using_random_quantum_circuits)

[researchgate.net](https://www.researchgate.net/publication/377136235_Random_projection_using_random_quantum_circuits)

[(PDF) Random projection using random quantum circuits - ResearchGate](https://www.researchgate.net/publication/377136235_Random_projection_using_random_quantum_circuits)

[Opens in a new window](https://www.researchgate.net/publication/377136235_Random_projection_using_random_quantum_circuits)

[](https://developers.google.com/machine-learning/glossary/rl?hl=he)

[developers.google.com](https://developers.google.com/machine-learning/glossary/rl?hl=he)

[מילון מונחים ללמידת מכונה: למידת חיזוק | Machine Learning - Google for Developers](https://developers.google.com/machine-learning/glossary/rl?hl=he)

[Opens in a new window](https://developers.google.com/machine-learning/glossary/rl?hl=he)

[](https://he.wikipedia.org/wiki/%D7%9C%D7%9E%D7%99%D7%93%D7%AA_%D7%97%D7%99%D7%96%D7%95%D7%A7)

[he.wikipedia.org](https://he.wikipedia.org/wiki/%D7%9C%D7%9E%D7%99%D7%93%D7%AA_%D7%97%D7%99%D7%96%D7%95%D7%A7)

[למידת חיזוק - ויקיפדיה](https://he.wikipedia.org/wiki/%D7%9C%D7%9E%D7%99%D7%93%D7%AA_%D7%97%D7%99%D7%96%D7%95%D7%A7)

[Opens in a new window](https://he.wikipedia.org/wiki/%D7%9C%D7%9E%D7%99%D7%93%D7%AA_%D7%97%D7%99%D7%96%D7%95%D7%A7)

[](https://pettingzoo.farama.org/)

[pettingzoo.farama.org](https://pettingzoo.farama.org/)

[PettingZoo Documentation](https://pettingzoo.farama.org/)

[Opens in a new window](https://pettingzoo.farama.org/)

[](https://paperswithcode.com/task/reinforcement-learning-1?page=11&q=)

[paperswithcode.com](https://paperswithcode.com/task/reinforcement-learning-1?page=11&q=)

[Reinforcement Learning (RL) | Papers With Code](https://paperswithcode.com/task/reinforcement-learning-1?page=11&q=)

[Opens in a new window](https://paperswithcode.com/task/reinforcement-learning-1?page=11&q=)

[](https://www.reddit.com/r/algotrading/comments/1hrwlna/stock_market_prediction_with_deep_reinforcement/)

[reddit.com](https://www.reddit.com/r/algotrading/comments/1hrwlna/stock_market_prediction_with_deep_reinforcement/)

[Stock Market Prediction with Deep Reinforcement Learning : r/algotrading - Reddit](https://www.reddit.com/r/algotrading/comments/1hrwlna/stock_market_prediction_with_deep_reinforcement/)

[Opens in a new window](https://www.reddit.com/r/algotrading/comments/1hrwlna/stock_market_prediction_with_deep_reinforcement/)

[](https://scholarworks.umass.edu/server/api/core/bitstreams/17fead7c-0656-4dc3-8e64-cb92bde9b7f5/content)

[scholarworks.umass.edu](https://scholarworks.umass.edu/server/api/core/bitstreams/17fead7c-0656-4dc3-8e64-cb92bde9b7f5/content)

[Electric Power Market Modeling with Multi-Agent Reinforcement Learning - ScholarWorks@UMass](https://scholarworks.umass.edu/server/api/core/bitstreams/17fead7c-0656-4dc3-8e64-cb92bde9b7f5/content)

[Opens in a new window](https://scholarworks.umass.edu/server/api/core/bitstreams/17fead7c-0656-4dc3-8e64-cb92bde9b7f5/content)

[](https://inspirehep.net/literature/2862748)

[inspirehep.net](https://inspirehep.net/literature/2862748)

[Quantum framework for Reinforcement Learning: integrating Markov Decision Process, quantum arithmetic, and trajectory search - Inspire HEP](https://inspirehep.net/literature/2862748)

[Opens in a new window](https://inspirehep.net/literature/2862748)

[](https://nautilustrader.io/docs/latest/concepts/backtesting/)

[nautilustrader.io](https://nautilustrader.io/docs/latest/concepts/backtesting/)

[Backtesting | NautilusTrader Documentation](https://nautilustrader.io/docs/latest/concepts/backtesting/)

[Opens in a new window](https://nautilustrader.io/docs/latest/concepts/backtesting/)

[](https://slogix.in/machine-learning/research-topics-in-deep-multiagent-reinforcement-learning/)

[slogix.in](https://slogix.in/machine-learning/research-topics-in-deep-multiagent-reinforcement-learning/)

[Research Topics in Deep Multiagent Reinforcement Learning - S-Logix](https://slogix.in/machine-learning/research-topics-in-deep-multiagent-reinforcement-learning/)

[Opens in a new window](https://slogix.in/machine-learning/research-topics-in-deep-multiagent-reinforcement-learning/)

[](https://arxiv.org/abs/2407.10403)

[arxiv.org](https://arxiv.org/abs/2407.10403)

[[2407.10403] Cooperative Reward Shaping for Multi-Agent Pathfinding - arXiv](https://arxiv.org/abs/2407.10403)

[Opens in a new window](https://arxiv.org/abs/2407.10403)

[](https://arxiv.org/html/2410.14927v1)

[arxiv.org](https://arxiv.org/html/2410.14927v1)

[Hierarchical Reinforced Trader (HRT): A Bi-Level Approach for Optimizing Stock Selection and Execution - arXiv](https://arxiv.org/html/2410.14927v1)

[Opens in a new window](https://arxiv.org/html/2410.14927v1)

[](https://datafloq.com/read/understanding-multi-agent-reinforcement-learning-marl/)

[datafloq.com](https://datafloq.com/read/understanding-multi-agent-reinforcement-learning-marl/)

[Understanding Multi-Agent Reinforcement Learning (MARL) - Datafloq](https://datafloq.com/read/understanding-multi-agent-reinforcement-learning-marl/)

[Opens in a new window](https://datafloq.com/read/understanding-multi-agent-reinforcement-learning-marl/)

[](https://www.numberanalytics.com/blog/rl-markets-math-guide)

[numberanalytics.com](https://www.numberanalytics.com/blog/rl-markets-math-guide)

[RL in Markets: A Neat Math How-To Guide - Number Analytics](https://www.numberanalytics.com/blog/rl-markets-math-guide)

[Opens in a new window](https://www.numberanalytics.com/blog/rl-markets-math-guide)

[](https://www.researchgate.net/publication/369414060_Strategic_Trading_in_Quantitative_Markets_through_Multi-Agent_Reinforcement_Learning)

[researchgate.net](https://www.researchgate.net/publication/369414060_Strategic_Trading_in_Quantitative_Markets_through_Multi-Agent_Reinforcement_Learning)

[Strategic Trading in Quantitative Markets through Multi-Agent Reinforcement Learning](https://www.researchgate.net/publication/369414060_Strategic_Trading_in_Quantitative_Markets_through_Multi-Agent_Reinforcement_Learning)

[Opens in a new window](https://www.researchgate.net/publication/369414060_Strategic_Trading_in_Quantitative_Markets_through_Multi-Agent_Reinforcement_Learning)

[](https://www.neuralnet.ai/designing-quantum-circuits-with-reinforcement-learning/)

[neuralnet.ai](https://www.neuralnet.ai/designing-quantum-circuits-with-reinforcement-learning/)

[Designing Quantum Circuits with Reinforcement Learning - NeuralNet.ai](https://www.neuralnet.ai/designing-quantum-circuits-with-reinforcement-learning/)

[Opens in a new window](https://www.neuralnet.ai/designing-quantum-circuits-with-reinforcement-learning/)

[](https://github.com/MultiAgentRL-trading/MARL)

[github.com](https://github.com/MultiAgentRL-trading/MARL)

[MultiAgentRL-trading/MARL - GitHub](https://github.com/MultiAgentRL-trading/MARL)

[Opens in a new window](https://github.com/MultiAgentRL-trading/MARL)

[](https://arxiv.org/abs/2412.20138)

[arxiv.org](https://arxiv.org/abs/2412.20138)

[[2412.20138] TradingAgents: Multi-Agents LLM Financial Trading Framework - arXiv](https://arxiv.org/abs/2412.20138)

[Opens in a new window](https://arxiv.org/abs/2412.20138)

[](https://arxiv.org/abs/2201.07065)

[arxiv.org](https://arxiv.org/abs/2201.07065)

[arxiv.org](https://arxiv.org/abs/2201.07065)

[Opens in a new window](https://arxiv.org/abs/2201.07065)

[](https://arxiv.org/abs/1905.05451)

[arxiv.org](https://arxiv.org/abs/1905.05451)

[arxiv.org](https://arxiv.org/abs/1905.05451)

[Opens in a new window](https://arxiv.org/abs/1905.05451)

[](https://arxiv.org/abs/2110.08448)

[arxiv.org](https://arxiv.org/abs/2110.08448)

[arxiv.org](https://arxiv.org/abs/2110.08448)

[Opens in a new window](https://arxiv.org/abs/2110.08448)

[](https://arxiv.org/abs/2305.16709)

[arxiv.org](https://arxiv.org/abs/2305.16709)

[arxiv.org](https://arxiv.org/abs/2305.16709)

[Opens in a new window](https://arxiv.org/abs/2305.16709)

[](https://arxiv.org/html/2502.05453v1)

[arxiv.org](https://arxiv.org/html/2502.05453v1)

[LLM-Powered Decentralized Generative Agents with Adaptive Hierarchical Knowledge Graph for Cooperative Planning - arXiv](https://arxiv.org/html/2502.05453v1)

[Opens in a new window](https://arxiv.org/html/2502.05453v1)

[](https://www.activeloop.ai/resources/glossary/multi-agent-reinforcement-learning-marl/)

[activeloop.ai](https://www.activeloop.ai/resources/glossary/multi-agent-reinforcement-learning-marl/)

[What is Multi-Agent Reinforcement Learning (MARL) - Activeloop](https://www.activeloop.ai/resources/glossary/multi-agent-reinforcement-learning-marl/)

[Opens in a new window](https://www.activeloop.ai/resources/glossary/multi-agent-reinforcement-learning-marl/)

[](https://github.com/AI4Finance-Foundation/FinRL-Tutorials/blob/master/Awesome_FinRL.md)

[github.com](https://github.com/AI4Finance-Foundation/FinRL-Tutorials/blob/master/Awesome_FinRL.md)

[FinRL-Tutorials/Awesome\_FinRL.md at master - GitHub](https://github.com/AI4Finance-Foundation/FinRL-Tutorials/blob/master/Awesome_FinRL.md)

[Opens in a new window](https://github.com/AI4Finance-Foundation/FinRL-Tutorials/blob/master/Awesome_FinRL.md)

[](https://arxiv.org/html/2503.04262)

[arxiv.org](https://arxiv.org/html/2503.04262)

[Guidelines for Applying RL and MARL in Cybersecurity Applications - arXiv](https://arxiv.org/html/2503.04262)

[Opens in a new window](https://arxiv.org/html/2503.04262)

[](https://relevanceai.com/agent-templates-tasks/data-aggregation-rules)

[relevanceai.com](https://relevanceai.com/agent-templates-tasks/data-aggregation-rules)

[Data Aggregation Rules AI Agents - Relevance AI](https://relevanceai.com/agent-templates-tasks/data-aggregation-rules)

[Opens in a new window](https://relevanceai.com/agent-templates-tasks/data-aggregation-rules)

[](https://arxiv.org/html/2503.13415v1)

[arxiv.org](https://arxiv.org/html/2503.13415v1)

[A Comprehensive Survey on Multi-Agent Cooperative Decision-Making: Scenarios, Approaches, Challenges and Perspectives - arXiv](https://arxiv.org/html/2503.13415v1)

[Opens in a new window](https://arxiv.org/html/2503.13415v1)

[](https://proceedings.neurips.cc/paper/2021/file/20aee3a5f4643755a79ee5f6a73050ac-Paper.pdf)

[proceedings.neurips.cc](https://proceedings.neurips.cc/paper/2021/file/20aee3a5f4643755a79ee5f6a73050ac-Paper.pdf)

[Celebrating Diversity in Shared Multi-Agent Reinforcement Learning](https://proceedings.neurips.cc/paper/2021/file/20aee3a5f4643755a79ee5f6a73050ac-Paper.pdf)

[Opens in a new window](https://proceedings.neurips.cc/paper/2021/file/20aee3a5f4643755a79ee5f6a73050ac-Paper.pdf)

[](https://philarchive.org/archive/DHRQMM)

[philarchive.org](https://philarchive.org/archive/DHRQMM)

[Merging quantum computing with next-gen AI - PhilArchive](https://philarchive.org/archive/DHRQMM)

[Opens in a new window](https://philarchive.org/archive/DHRQMM)

[](https://www.sevenmentor.com/power-of-python-in-data-analytics)

[sevenmentor.com](https://www.sevenmentor.com/power-of-python-in-data-analytics)

[Power of Python in Data Analytics | SevenMentor](https://www.sevenmentor.com/power-of-python-in-data-analytics)

[Opens in a new window](https://www.sevenmentor.com/power-of-python-in-data-analytics)

[](https://niall.phd/pdfs/publications/terry2020petting.pdf)

[niall.phd](https://niall.phd/pdfs/publications/terry2020petting.pdf)

[PettingZoo: Gym for Multi-Agent Reinforcement Learning - Niall L. Williams](https://niall.phd/pdfs/publications/terry2020petting.pdf)

[Opens in a new window](https://niall.phd/pdfs/publications/terry2020petting.pdf)

[](https://arxiv.org/html/2411.11099v1)

[arxiv.org](https://arxiv.org/html/2411.11099v1)

[Mitigating Relative Over-Generalization in Multi-Agent Reinforcement Learning - arXiv](https://arxiv.org/html/2411.11099v1)

[Opens in a new window](https://arxiv.org/html/2411.11099v1)

[](https://unstructured.io/blog/building-an-end-to-end-data-pipeline-with-custom-ner-on-unstructured-using-mcp)

[unstructured.io](https://unstructured.io/blog/building-an-end-to-end-data-pipeline-with-custom-ner-on-unstructured-using-mcp)

[Building an End-to-End Data Pipeline with Custom NER on Unstructured using MCP](https://unstructured.io/blog/building-an-end-to-end-data-pipeline-with-custom-ner-on-unstructured-using-mcp)

[Opens in a new window](https://unstructured.io/blog/building-an-end-to-end-data-pipeline-with-custom-ner-on-unstructured-using-mcp)

[](https://estuary.dev/blog/finance-data-modernization-real-time-sync/)

[estuary.dev](https://estuary.dev/blog/finance-data-modernization-real-time-sync/)

[Modernize Financial Data Infrastructure with Real-Time Sync—No Rewrites Needed](https://estuary.dev/blog/finance-data-modernization-real-time-sync/)

[Opens in a new window](https://estuary.dev/blog/finance-data-modernization-real-time-sync/)

[](https://arxiv.org/html/2502.08985v1)

[arxiv.org](https://arxiv.org/html/2502.08985v1)

[Few is More: Task-Efficient Skill-Discovery for Multi-Task Offline Multi-Agent Reinforcement Learning - arXiv](https://arxiv.org/html/2502.08985v1)

[Opens in a new window](https://arxiv.org/html/2502.08985v1)

[](https://github.com/Jaroan/Fair-MARL)

[github.com](https://github.com/Jaroan/Fair-MARL)

[Jaroan/Fair-MARL: Cooperation and Fairness in Multi-Agent Reinforcement Learning - GitHub](https://github.com/Jaroan/Fair-MARL)

[Opens in a new window](https://github.com/Jaroan/Fair-MARL)

[](https://arxiv.org/html/2504.14520v1)

[arxiv.org](https://arxiv.org/html/2504.14520v1)

[Meta‑Thinking in LLMs via Multi‑Agent Reinforcement Learning: A Survey - arXiv](https://arxiv.org/html/2504.14520v1)

[Opens in a new window](https://arxiv.org/html/2504.14520v1)

[](https://fairvote.org/report/fairvotes-position-on-fusion-voting/)

[fairvote.org](https://fairvote.org/report/fairvotes-position-on-fusion-voting/)

[FairVote's Position on Fusion Voting](https://fairvote.org/report/fairvotes-position-on-fusion-voting/)

[Opens in a new window](https://fairvote.org/report/fairvotes-position-on-fusion-voting/)

[](https://www.researchgate.net/publication/385350757_Quantum-Inspired_Algorithms_for_AI_and_Machine_Learning)

[researchgate.net](https://www.researchgate.net/publication/385350757_Quantum-Inspired_Algorithms_for_AI_and_Machine_Learning)

[Quantum-Inspired Algorithms for AI and Machine Learning - ResearchGate](https://www.researchgate.net/publication/385350757_Quantum-Inspired_Algorithms_for_AI_and_Machine_Learning)

[Opens in a new window](https://www.researchgate.net/publication/385350757_Quantum-Inspired_Algorithms_for_AI_and_Machine_Learning)

[](https://www.globenewswire.com/news-release/2025/05/16/3083101/0/en/MicroCloud-Hologram-Inc-Develops-Neural-Network-Based-Quantum-Assisted-Unsupervised-Data-Clustering-Technology.html)

[globenewswire.com](https://www.globenewswire.com/news-release/2025/05/16/3083101/0/en/MicroCloud-Hologram-Inc-Develops-Neural-Network-Based-Quantum-Assisted-Unsupervised-Data-Clustering-Technology.html)

[MicroCloud Hologram Inc. Develops Neural Network-Based Quantum-Assisted Unsupervised Data Clustering Technology - GlobeNewswire](https://www.globenewswire.com/news-release/2025/05/16/3083101/0/en/MicroCloud-Hologram-Inc-Develops-Neural-Network-Based-Quantum-Assisted-Unsupervised-Data-Clustering-Technology.html)

[Opens in a new window](https://www.globenewswire.com/news-release/2025/05/16/3083101/0/en/MicroCloud-Hologram-Inc-Develops-Neural-Network-Based-Quantum-Assisted-Unsupervised-Data-Clustering-Technology.html)

[](https://www.researchgate.net/publication/389446669_Solving_Action_Semantic_Conflict_in_Physically_Heterogeneous_Multi-Agent_Reinforcement_Learning_with_Generalized_Action-Prediction_Optimization)

[researchgate.net](https://www.researchgate.net/publication/389446669_Solving_Action_Semantic_Conflict_in_Physically_Heterogeneous_Multi-Agent_Reinforcement_Learning_with_Generalized_Action-Prediction_Optimization)

[(PDF) Solving Action Semantic Conflict in Physically Heterogeneous Multi-Agent Reinforcement Learning with Generalized Action-Prediction Optimization - ResearchGate](https://www.researchgate.net/publication/389446669_Solving_Action_Semantic_Conflict_in_Physically_Heterogeneous_Multi-Agent_Reinforcement_Learning_with_Generalized_Action-Prediction_Optimization)

[Opens in a new window](https://www.researchgate.net/publication/389446669_Solving_Action_Semantic_Conflict_in_Physically_Heterogeneous_Multi-Agent_Reinforcement_Learning_with_Generalized_Action-Prediction_Optimization)

[](https://arxiv.org/html/2502.05812v1)

[arxiv.org](https://arxiv.org/html/2502.05812v1)

[Multi-Agent Reinforcement Learning in Wireless Distributed Networks for 6G - arXiv](https://arxiv.org/html/2502.05812v1)

[Opens in a new window](https://arxiv.org/html/2502.05812v1)

[](https://arxiv.org/pdf/1905.05451)

[arxiv.org](https://arxiv.org/pdf/1905.05451)

[Moment-Based Variational Inference for Markov Jump Processes - arXiv](https://arxiv.org/pdf/1905.05451)

[Opens in a new window](https://arxiv.org/pdf/1905.05451)

[](https://arxiv.org/pdf/2309.01022)

[arxiv.org](https://arxiv.org/pdf/2309.01022)

[Overleaf Example - arXiv](https://arxiv.org/pdf/2309.01022)

[Opens in a new window](https://arxiv.org/pdf/2309.01022)

[](https://arxiv.org/pdf/2304.05089)

[arxiv.org](https://arxiv.org/pdf/2304.05089)

[arXiv:2304.05089v1 [hep-ph] 11 Apr 2023](https://arxiv.org/pdf/2304.05089)

[Opens in a new window](https://arxiv.org/pdf/2304.05089)

[](https://arxiv.org/pdf/2110.08448)

[arxiv.org](https://arxiv.org/pdf/2110.08448)

[Deepest Cuts for Benders Decomposition - arXiv](https://arxiv.org/pdf/2110.08448)

[Opens in a new window](https://arxiv.org/pdf/2110.08448)

[](https://arxiv.org/pdf/2403.07433)

[arxiv.org](https://arxiv.org/pdf/2403.07433)

[Baryonic Vortex Phase and Magnetic Field Generation in QCD with Isospin and Baryon Chemical Potentials - arXiv](https://arxiv.org/pdf/2403.07433)

[Opens in a new window](https://arxiv.org/pdf/2403.07433)

[](https://www.mdpi.com/2412-3811/10/5/112)

[mdpi.com](https://www.mdpi.com/2412-3811/10/5/112)

[A Deep Reinforcement Learning Framework for Last-Mile Delivery with Public Transport and Traffic-Aware Integration: A Case Study in Casablanca - MDPI](https://www.mdpi.com/2412-3811/10/5/112)

[Opens in a new window](https://www.mdpi.com/2412-3811/10/5/112)

[](https://arxiv.org/html/2412.20138)

[arxiv.org](https://arxiv.org/html/2412.20138)

[TradingAgents: Multi-Agents LLM Financial Trading Framework - arXiv](https://arxiv.org/html/2412.20138)

[Opens in a new window](https://arxiv.org/html/2412.20138)

[](https://www.arxiv.org/pdf/2505.03949)

[arxiv.org](https://www.arxiv.org/pdf/2505.03949)

[Deep Q-Network (DQN) multi-agent reinforcement learning (MARL) for Stock Trading - arXiv](https://www.arxiv.org/pdf/2505.03949)

[Opens in a new window](https://www.arxiv.org/pdf/2505.03949)

[](https://www.annualreviews.org/content/journals/10.1146/annurev-statistics-112723-034423)

[annualreviews.org](https://www.annualreviews.org/content/journals/10.1146/annurev-statistics-112723-034423)

[A Review of Reinforcement Learning in Financial Applications](https://www.annualreviews.org/content/journals/10.1146/annurev-statistics-112723-034423)

[Opens in a new window](https://www.annualreviews.org/content/journals/10.1146/annurev-statistics-112723-034423)

[](https://ojs.aaai.org/index.php/AAAI/article/view/34045/36200)

[ojs.aaai.org](https://ojs.aaai.org/index.php/AAAI/article/view/34045/36200)

[Logic-Q: Improving Deep Reinforcement Learning-based Quantitative Trading via Program Sketch-based Tuning - AAAI Publications](https://ojs.aaai.org/index.php/AAAI/article/view/34045/36200)

[Opens in a new window](https://ojs.aaai.org/index.php/AAAI/article/view/34045/36200)

[](https://callforpapers.institutlouisbachelier.org/Papers/27aa3d21-bda0-4140-9b6e-cee2ab86b8f4.pdf)

[callforpapers.institutlouisbachelier.org](https://callforpapers.institutlouisbachelier.org/Papers/27aa3d21-bda0-4140-9b6e-cee2ab86b8f4.pdf)

[DATA TIME TRAVEL AND CONSISTENT MARKET MAKING: TAMING REINFORCEMENT LEARNING IN MULTI-AGENT SYSTEMS WITH ANONYMOUS DATA](https://callforpapers.institutlouisbachelier.org/Papers/27aa3d21-bda0-4140-9b6e-cee2ab86b8f4.pdf)

[Opens in a new window](https://callforpapers.institutlouisbachelier.org/Papers/27aa3d21-bda0-4140-9b6e-cee2ab86b8f4.pdf)

[](https://fastercapital.com/topics/strategies-for-trading-with-level-2-data.html)

[fastercapital.com](https://fastercapital.com/topics/strategies-for-trading-with-level-2-data.html)

[Strategies For Trading With Level 2 Data - FasterCapital](https://fastercapital.com/topics/strategies-for-trading-with-level-2-data.html)

[Opens in a new window](https://fastercapital.com/topics/strategies-for-trading-with-level-2-data.html)

[](https://arxiv.org/pdf/2403.16801)

[arxiv.org](https://arxiv.org/pdf/2403.16801)

[arXiv:2403.16801v2 [hep-th] 2 Sep 2024](https://arxiv.org/pdf/2403.16801)

[Opens in a new window](https://arxiv.org/pdf/2403.16801)

[](https://arxiv.org/pdf/2305.16709)

[arxiv.org](https://arxiv.org/pdf/2305.16709)

[Zero-modes in magnetized T6/ZN orbifold models through Sp(6,Z) modular symmetry - arXiv](https://arxiv.org/pdf/2305.16709)

[Opens in a new window](https://arxiv.org/pdf/2305.16709)

[](https://www.researchgate.net/publication/388685307_CH-MARL_Constrained_Hierarchical_Multiagent_Reinforcement_Learning_for_Sustainable_Maritime_Logistics)

[researchgate.net](https://www.researchgate.net/publication/388685307_CH-MARL_Constrained_Hierarchical_Multiagent_Reinforcement_Learning_for_Sustainable_Maritime_Logistics)

[CH-MARL: Constrained Hierarchical Multiagent Reinforcement Learning for Sustainable Maritime Logistics - ResearchGate](https://www.researchgate.net/publication/388685307_CH-MARL_Constrained_Hierarchical_Multiagent_Reinforcement_Learning_for_Sustainable_Maritime_Logistics)

[Opens in a new window](https://www.researchgate.net/publication/388685307_CH-MARL_Constrained_Hierarchical_Multiagent_Reinforcement_Learning_for_Sustainable_Maritime_Logistics)

[](https://www.youtube.com/watch?v=7EXNr5GlxTs)

[youtube.com](https://www.youtube.com/watch?v=7EXNr5GlxTs)

[Trading Using LLM | Generative AI & Sentiment Analysis for Finance | Webinar - YouTube](https://www.youtube.com/watch?v=7EXNr5GlxTs)

[Opens in a new window](https://www.youtube.com/watch?v=7EXNr5GlxTs)

Thoughts

מושגי יסוד ומשאבים

התחלתי במיפוי הרעיונות המרכזיים בפרויקט, כמו למידת חיזוק מרובת סוכנים (MARL) למסחר, נתוני שוק Level 2 וזיהוי משטרי שוק. בחנתי את כל הקישורים שסיפקת כדי להבין אילו כלים ושיטות קיימים שיכולים לעזור לי. שמתי לב במיוחד למאמרים שעוסקים ב-MARL, זיהוי משטרים ויישומים של למידה עמוקה במסחר פיננסי.

גבי ההנדסה של הנתונים - חשוב מאוד שנבין שהמערכת בעצם תעבוד ישירות על מסחר בזמן אמת עם כסף חי - כך שהמטריצות יהיו חייבות להתעדכן בצורה כזו או אחרת - אגב למה כל מצטריצה לא יכולה לכלול את כל סוגי המאפיינים הרי זה הכי הגיוני - כל יום מסחר ביחוד על מערכות מסחר תוך יומיות אנחנו צפויים לחוות את כל אלה ביום מסחר אחד - או אפילו בסשן אחד - מה עחושים כדי לפתור את הבעיה הזאת? אני חושב שנוריד ונוותר על שימוש בנתונים משלב 2 - הם יוצרים המון רעש - אני חושב שאם לחלק את הדאטה למטריצות וגם להוסיף את המנוע לזיהוי המשטר יכול להיות די והותר - מאחר והזיהוי יתקיים על נפח זמן של 30 דק לכל בר - הוא יהיה מאוד מקיף ולהכניס למטריצה הרבה רעש יכול להיות אובדני - אלא אם משתמשים בPCA/SVD כדי להוריד ממדים זה יהיה מאוד מעניין - אבל עוד בעיה שהנתונים הם אינם מחושבים באותה צורה וכך שנרמול הנתוני כדי לבצע PCA /SVD איכותי יכול להיווצר לנו בעיה אחרת שאנחנו מורידים ממדים מאוד חשובים וצמשאירים את הלא רלוונטים - אכן שאלה מאוד כבדה של מה ניתן לעשות - יכול להיות שיש פתרון - חייב לבדוק את הנושא - אגב אני מציע רעיון - המנוע של זיהוי המשטר מוציא נתונים - למה לא להתמשמש בנתונים שכבר יש לנו והם פחות רעשנים שיעזור לסוכן להבין יותר את המימירות של השוק ולא לרדת לנפחי זמן של פחות משניה וזה לדעתי יעזור להפיק ערך רב יותר עבור הסוכן אם נוריד שיחמוש בשללב 2 ונוסיף נתונים מהמנוע לזיהוי משטרי שוק - האתגר המרכזי בשיתוף פעולה, כאשר סוכנים פועלים על סמך תצפיות חלקיות ואקראיות, הוא כיצד להבטיח שפעולותיהם, הנגזרות מאותה מדיניות, יהיו אכן מתואמות ולא סותרות ברמה הגלובלית. ייתכן שהמדיניות המשותפת תצטרך להיות מותנית גם על רכיב מצב גלובלי כלשהו (מעבר לפרמטרי המשטר), או שמבנה התגמול יצטרך להעניש בחריפות חוסר תיאום. - נקודה זו שנוקבת בבעיה המרכזית בנוגע לשת"פ בין כלל הסוכנים הניה קריטיצ וקוהרנטית לפתרון - איך משערכים מה יכול להיות או ללהוות יציבות מסויימת לפרמטר - 3.2.1. ייצוג ארבע הסינרגיות (לונג/שורט) במדיניות ה-RL

יש להגדיר במדויק את "ארבע הסינרגיות" ואת התנאים להפעלתן. רשת המדיניות של סוכני ה-RL צריכה להיות מסוגלת לייצג ולבצע את הסינרגיות הללו. קיימות מספר אפשרויות לייצוג כזה: - אין צורך לתת לסוכנים להחליט איזו סינרגיה לבחור זה לא רלוונטי - בעצם הסינרגיה היא בין שלושת האינדיקטורים עליהם בנויה כל האסטרטגיה הזו - כל סינרגיה מוצגת באופן שונה כי יכול להיות מצב שעכשיו יש לנו אפשרות כניסה עם סינרגיה 2 ויום למחורת עם סינרגיה 4 לצורך העניין - מה שאני בא להסביר כן זה שהשוק מייצר את המצבים - האינדיקטורים מייצרים סיגנל ושלושת הסיגנלים ביחד יוצרים את האפשרו ת להיכנס לעסקה - יכול להיווצר ארבעה אפשויות שונה לכניסה לעסקה וכל אפשרות מוצגת כסינרגיה כי בעצם שלושת האינדיקטורים ביחד בסופו של יום נותנים לנו סיגנל אחד שהוא סיגנל הכניסה - אז הסוכן לא יצטרך להבין איזה סינרגיה לבחור השוק בוחר בשבילו - הסוכן יצטרך לדעת איך נכון להיכנס לעסקה ואיך נכון לנהל את הסיכונים - למשל הסוכן יכול להתשתמש בצורות הרמוניות - ובעצם כל פעם שיש סיגנל לנקוב נקודה ובסוף שיש לך את שלושת הסיגנלים ביחד יש לנו משולש בעצם עם זוויות וצלעות וגישה זו של לקחת את שלושת הסיגנלים ולא להציג אותם רק כדיגנל אלא גם כצורה גיאומרטים (במקרה שלנו זה משולש) יכול לתת המון לסוכנים להבין אילו כניסות הם נכונות ואלו כניסות לא נכונות ועל ביס הצורה הגיאומטרית שנוצרה בעצם - יש לנו אפשרות לקשר את זה לניהול סיכונים ולהגיד אוקיי בהסתכלות אחורנית צורה עם התכונות האלה הייתה טובה יותר אפשר להיכנס עם םוזיציה גדולה יותר ממש כמו שבני אדם רגילים עושים פשוט פה מפשטים את זה בצורה האפשרית שיש כדי לאפשר לסוכנים להבין הכל מהכל! אז לסיכום השוק זז - האינדיקטורים מוציאים ססיגנל - כשיש את שלושת הסיגנים ביחד -אחת מהסינרגיות תהיהנ חייבת להתממש - ואך ורק במצב כזה אנחנו יכולים להגיד לסוכן הנה זה המצב - איך נכנסים כאן? אגב עוד פרט מידע סופר חשוב - השוק זז בין תחנות וזה לא משנה איזה נכס כולם זזים בין תחנות - אז המחשבה שלי היא שאנחנו מבצעים איתה מסחר ידנני היא לקחת את פרופיל השוק market profile ולסמן LVN (low volume nodes) ובעצם כל פעם שהשוק מגיע לתחנה - פוטנציאל לעסקה מצליחה לשני יהכיוונים גם לונג וגם שורט עולים כי שאני יודע שלאזור הזה המחיר רצה להגיע - וכך גם ניתן להעריך את חוזק התחנה - אם התחנה הזאת יצרה תמיכה והתנגדות בכמה אזורים נדע שהי מאוד חזקה ואולי שם נכנס עם פוזיצייית ענק ולעומת זאת אם זאת תחנה חדשה נכנס עם פוזיציה קטנה יותר - הסוכן יכול להעריך את עווצמת התחנה על בסיס העבר כי השוק זז ע"פ זה - הם יוצר תחנות חדשות וחוזר אליהם אז זה בעצםם מחייב אותנו גם להעיריך סטטיסטית כל תחנה והשימוש גם בצורות הרמוניות שמקשרות את הסיגנלים ביחד ויוצאת לנו צורה גיאומטרית עם זויות וצלעטות שניתן להגיד לסוכן הזה תראה זה המצב הצורה הזאתיציבה ותואמת לסינרגיה - כנס עם פוזיציה גדולה - וכמובן שימוש בתחנות שזוהו באמצעות LVN של Market profile יכול להיות שהסוכנים יבינו את האיכותבאזוורים האלה ויכנסו רק שם וככל שיש תחנות חזקות יותר כך הכניסה תהיה צמסיבית יותר וניהול הסיכונים יהיה בהתאמה לצורה ההרמונית שמלכדת את שלושת הסיגנלים ביחד - אני שוב מציין בהתאם לסעיף 3.2.2 המערכתתסחר עם כסף חי! יש לשים לזה לב - שיטות הטבעה (Embedding): שימוש ברשתות נוירונים (למשל, רשתות קונבולוציה או רשתות רקורנטיות) כדי ללמוד ייצוג דחוס ואינפורמטיבי של המקטע Mi

​.האם זה יכול להיות פתרון לבעיית האיפיון שלנו של הפרמטרים? אני לא יודע אם נצטרך נתונים שלב 2 אבל בכל אופן אם כן נצטרך ואם לא שיטה זא יכולה לסגר את כל הדחיסה של הנצתוניםן למצטרימה ולאפשר לסוכן סביבה הרבה יותר רציפה איכותית ויציבה - מאוד אהבי בוא נחקור שיטות ליישום בהתחשב בעבודה של מערכות מסגננון כזה - מחקר 1 מציע הפשטת תצפיות על בסיס יחסים מרחביים לתוך ייצוג מבוסס-גרף (MARC), רעיון שאף על פי שהוא מיועד למשימות מרחביות, מדגיש את החשיבות של יצירת ייצוג מובנה ומופשט מהתצפית הגולמית. בספריות MARL כמו Ray RLlib, תצפיות מוחזרות כמילונים הממפים מזהי סוכנים לתצפיות האישיות שלהם 2 , מה שמתאים למבנה המוצע - לא הבנתי - תסביר לי בבקשה מה זה אומר איך זה בא לידי ביטוי ולמה זה דווקא המבנה המוצע - אגב - טווחי הזמן של האסטרטגיה הינם 30 דק וגם זיהוי התחנות יתבצע שם - והטווח זמן השני של ביצוע העסה יהיה ב5 דק - ששם יתבצע התחיה בעצם של הצורה ההרמונית ולמרות שיש לנו שני אינדיקטורים במיוצגים בטווח הזמן של 30 דק הם ייצרו סיגנל והוא יוצג והצורה ההרמונית בעצם תצוייר על וןןח הזמן של 5 דק - אני כמובן אתאר פה את האסטרטגיה לעומק ושמם יהיה קל מאוד לחבר הכל! - אנילא הבנתי באיזה סיפרייה נכון לממש את MARL - זה עוד שלב מאוחר -ולפני כן יש לנו מימוש של המנווע לזיהוי מצבי שוק והמערכת שמחקלת את הנתונים בצורה רנדומלית למטריצות והמטריצות עצמן מיוצגות כווקטור שיאפשר לסוכנים סביבת אימון טובה יותר -צריך לבדוק אם יש מערכת כזאת - אם אין אז צריך להבין איך לבנות אחת כזאת ולאחד אותה עם המנוע ךזיהוי מצבי שוק באופן רובסטי - גם לנתוני עבר וגם לנתוני זמן אמת! - האתגר המרכזי בתכנון מערכת ה-MARL נובע מהשילוב הייחודי של מדיניות משותפת עם קלט נתונים אקראי ומבוזר. אם המדיניות המשותפת (הכוללת 4 סינרגיות) מורכבת, ומקטעי הנתונים Mi

​ נבדלים מאוד זה מזה או דלים במידע, הסוכנים עלולים ליישם את המדיניות באופן שגוי או לא עקבי. הדבר עלול להוביל לפעולות תת-אופטימליות או אף סותרות, למרות הכוונה לשיתוף פעולה. "שיתוף הפעולה המושלם" במקרה זה עשוי שלא לנבוע מלמידה של הסוכנים כיצד לשתף פעולה, אלא מהיכולת של המערכת להבטיח שפעולותיהם, הנגזרות מאותה מדיניות אך מיושמות על נתונים שונים, יהיו קוהרנטיות ברמה הגלובלית. הדבר מצביע על כך שהצלחת המערכת תלויה באופן קריטי ביכולת של המדיניות המשותפת להתמודד עם מגוון רחב של תשומות ממקטעי הנתונים השונים, ובכך שהסכום של הפעולות הלוקאליות-אופטימליות (הנגזרות מהמדיניות המשותפת על נתונים מקומיים) אכן יתקרב לאסטרטגיה גלובלית-אופטימלית. - בעצם רוב העבודה העיקרית היא לא המערכת MARL היא בעצם בניית מערכת שתדע לקחת נתונים ולסדר אותם נכון! זה הכי חשוב כדי להבטיח את הצלחתה של המערכת! - לגבי הסיפור של הסופר פוזיציה - ומה להטמיע - אנחנו נריץ מחקר מעמיק על זה לאחר ונבין מה הכי נכון יהיה לעשות בהסתכלםת עלהמערכת ולראות מה היכולות שלה וגם מה המוגבלות שלה וע"פפ עובדות אלו ניתן להעריך מה נעשה - 5.1. הגדרת קריטריונים להצלחה ויעדי רווח והפסד

הגדרת קריטריונים ברורים להצלחה היא צעד ראשון והכרחי. בעוד שהמשתמש ציין יעד ספציפי של "כמה האסטרטגיה צריכה להרוויח בשבועיים", חשוב להבין כי התמקדות צרה ביעד רווח והפסד (P&L) קצר טווח עלולה להוביל להתאמת יתר (overfitting). לכן, יש לאמץ מגוון רחב יותר של מדדי ביצועים:

מדדי תשואה מותאמת לסיכון: יחס שארפ (Sharpe Ratio), יחס סורטינו (Sortino Ratio), יחס קלמאר (Calmar Ratio). מדדים אלו מעריכים את התשואה ביחס לרמת הסיכון שנלקחה.

מדדי רווחיות: תשואה מצטברת (Cumulative Return), פקטור רווח (Profit Factor – סך הרווח חלקי סך ההפסד).

מדדי סיכון: משיכה מקסימלית (Maximum Drawdown – MDD), תנודתיות התשואות.

מדדי יעילות מסחר: אחוז עסקאות רווחיות (Win Rate), תוחלת רווח לעסקה (Expectancy).

הצלחת המערכת צריכה להימדד גם ביכולתה לשמור על ביצועים עקביים וחזקים על פני משטרי שוק שונים (כפי שזוהו על ידי מנוע ה-MMD) ועל פני מכשירים פיננסיים שונים. היעד של "רווח תוך שבועיים" יכול לשמש כנקודת ייחוס להערכה, אך אין לבצע אופטימיזציה ישירה ובלעדית אליו. - זה לגמרי יבוצע כפי שאתה מתאר - אנחנו נמדוד את כל הפרמטירים האלה ונדווא יציבות של מערכת - לגבי בעית הcredit assingment אני מעריך שזה שלב שהוא נבנה לאחר שיש הטמעה של סופרפוזיציה על המערכת נכון? לא הבנתי מתי שלב הבקטסט נכנס לתמונה - אחרי שהמערכת מוכנה? ואיך בדיוק זה אמור לעבוד בפועל - כאילו מה הסיפור עם הבקטסט ואיך מיישמים אתזה נכון ובודקים את המערכת - ואיך אימון באמצעות בק טסט קשור?רגולריזציה: הוספת רכיבי רגולריזציה לפונקציית המטרה של אלגוריתם ה-RL (למשל, L1/L2 על משקולות הרשת, עונש על מורכבות המדיניות). איך יישום של רגולריזציה צריך להתחבר? 6.3. אגרגציה ומיזוג מדיניות מתשומות/תצפיות הטרוגניות של סוכנים

אף על פי שכל הסוכנים חולקים את אותה מדיניות, הם מקבלים תשומות שונות (מטריצות Mi

​ שונות). אם הפלט של כל סוכן (למשל, הסתברויות לפעולה) צריך לעבור אגרגציה להחלטת צוות אחת (למשל, לצורך ניהול סיכונים כולל של התיק, או לביצוע פקודה נטו אחת בשוק), יש לדון כיצד לבצע אגרגציה זו.

הצבעת רוב או ממוצע משוקלל: ניתן לאסוף את המלצות הפעולה מכל הסוכנים ולבחור את הפעולה שקיבלה את רוב הקולות, או לחשב ממוצע משוקלל של הסתברויות הפעולה (כאשר המשקולות יכולות להתבסס על מדדי ביטחון של כל סוכן או על ביצועיו ההיסטוריים).

פרוטוקולי הצבעה מורכבים: מחקר דן במיזוג מדיניויות של סוכנים בעלי תגמולים שונים באמצעות מדדי תפוסת מצב-פעולה וכללי הצבעה כמו שיטת בורדה. רעיונות אלו עשויים להיות רלוונטיים אם ארבע הסינרגיות נתפסות כתת-מדיניויות שהסוכנים "מצביעים" עליהן.

מיזוג סוכנים מבוסס גיוון: גישת MARL-Focal , אף על פי שהיא מיועדת לסוכני LLM, מציעה רעיון של בחירת תת-קבוצה של סוכנים על בסיס יכולתם להשלים זה את זה, ושימוש בפתרון קונפליקטים ליצירת פלט מאוחד.

Value Decomposition Networks (VDN): כפי שהוזכר קודם, VDN ושיטות דומות (למשל, QMIX) מבצעות סוג של אגרגציית פונקציות ערך כדי להעריך את הפעולה המשותפת האופטימלית.

האינטראקציה בין חלוקת נתונים אקראית, מודעות למשטר מבוססת MMD, מדיניות משותפת מורכבת (4 סינרגיות), ורעיונות סופרפוזיציה לעמידות סוכנים, יוצרת מערכת שבה התנהגות אמרגנטית (emergent behavior) היא סבירה ביותר. התנהגות זו יכולה להיות מועילה (אסטרטגיות חדשניות ומסתגלות) או מזיקה (כשלים בלתי צפויים, פעולות כאוטיות). מפת הדרכים ליישום חייבת לכלול סימולציה מקיפה ובדיקות "ארגז חול" (sandboxing) כדי להבין תכונות אמרגנטיות אלו לפני כל פריסה חיה. שלב אגרגציית המדיניות (6.3) הופך לקריטי אם יש צורך לשלב את הפלטים האישיים של הסוכנים, במיוחד אם רעיון ה"סופרפוזיציה" מוביל לפלטים הסתברותיים או למספר פעולות מועמדות מכל סוכן. מאוד מאוד חשוב ליישום! מאוד! הנחיות למחקר עתידי:

הפיתוח והבדיקה של מערכת זו יניבו ככל הנראה תובנות רבות שיכולות להנחות מחקר עתידי:

טכניקות MMD מתקדמות יותר: חקירת גרסאות נוספות של MMD, אולי בשילוב עם שיטות למידה עמוקה נוספות, לזיהוי משטרים מדויק ומהיר יותר.

חלוקת נתונים דינמית או למידה: במקום חלוקה אקראית קבועה, ניתן לחקור שיטות בהן חלוקת הנתונים לסוכנים היא דינמית ומשתנה בהתאם לתנאי השוק, או אף נלמדת על ידי המערכת עצמה.

אבולוציה של סינרגיות: האפשרות שהמערכת תלמד או תפתח סינרגיות חדשות באופן אוטונומי, מעבר לארבע הסינרגיות המוגדרות מראש.

שילוב עמוק יותר של עקרונות בהשראת קוונטים: ככל שהבנתנו את האנלוגיות בין חישוב קוונטי ללמידה קלאסית תעמיק, וככל שחומרת ותוכנת החישוב הקלאסי יתפתחו, ייתכן שניתן יהיה לממש רעיונות "קוונטיים" מורכבים יותר באופן קלאסי לשיפור ביצועי סוכני RL.

פרשנות ושקיפות: פיתוח כלים וטכניקות להבנת תהליך קבלת ההחלטות של מערכת MARL מורכבת זו, במטרה להגביר את האמון בה ולהקל על איתור וטיפול בכשלים.

למידה מתמשכת והסתגלות ארוכת-טווח: חקירת ארכיטקטורות ומנגנוני למידה המאפשרים למערכת כולה להסתגל באופן רציף לשינויים ארוכי-טווח בדינמיקת השוק, מעבר לזיהוי משטרים טקטי.

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. ההתקדמות ההדרגתית, תוך התבססות על מחקר קיים וכלים מתקדמים, והתמודדות שיטתית עם האתגרים, היא המפתח להפיכת חזון זה למציאות. - אני רוצה להתייחס לשאלות להמשך המחקר - אם יש שיטה טובה יותר למצוא מצב שוק שיכול להביא תוצאות בצורה מהירה ואיכותית לא פחות מMMD ויודע כם להפחית שימוש בליבה החומרית של המחשב -זה מבורך ועדיף מצריך חקירה - חלוקת נתונים - אם יש אפרות לקחת את מצבי השוק ועל סמך זה לחלק נתונים בצורה שתתרום למערכת במקום לבנות מערכת סטטית שלא תזוז עם השוק אלא תתקע את הוסכנים ולא תמיד תהיה אדפטייבית - זה מדהים! איך מבצעים כזה דבר - צריל מערכת שתיקח את הנתונים ותחלק אותם באופן דינמי למצרטריצות ומשם לווקטורים כדי לאשפר לסוכנים גישה מהירה יותר - אבולוציה של סינרגיות היא מבורכת ועדיפה - אשמח לשמוע עוד על כך אבל לפני כן אני מצרף פה סיכום עלהאסטרטגיה שלי הסיכום הוא סיכום מעמיק זה חשוב שתכיר אותה לעומק - Analysis and Enhancement of a Multi-Indicator, Multi-Timeframe Algorithmic Trading Strategy

I. Introduction

Purpose: This report provides an in-depth analysis of a novel multi-indicator, multi-timeframe trading strategy operating on short timescales (1-5 minutes). The strategy leverages the Machine Learning Momentum Index (MLMI), Nadaraya-Watson Rational Quadratic Kernel Regression (NW-RQK), and Fair Value Gaps (FVG). The analysis addresses specific implementation challenges identified by the strategy's architect, namely dynamic adaptation to market regimes and the management of signal time validity. Furthermore, it evaluates the current exit mechanism and proposes refinements, culminating in the exploration of three advanced pathways aimed at achieving significantly enhanced ("super") robustness and performance.

Context: The proposed strategy operates within the demanding domain of short-term algorithmic trading. Modern approaches increasingly combine machine learning techniques, advanced statistical methods, and nuanced price action analysis to navigate complex market dynamics.1 However, high-frequency environments present inherent challenges, including pervasive market noise, non-stationarity (where statistical properties change over time), and the critical need for adaptive systems capable of responding to rapidly shifting conditions.3 The strategy's reliance on indicators across 1-minute and 5-minute charts necessitates careful consideration of these factors.

II. Strategy Analysis: MLMI, Nadaraya-Watson, and FVG Synergy

A. Component Breakdown

A comprehensive understanding of the strategy requires examining each of its constituent indicators.

Machine Learning Momentum Index (MLMI): The MLMI, developed by Zeiierman, functions as an oscillator that integrates traditional momentum concepts with machine learning.6 It calculates both a fast (5-period) and slow (20-period) Weighted Moving Average (WMA) of the Relative Strength Index (RSI) to gauge short-term and longer-term momentum. Crucially, it employs a k-Nearest Neighbors (k-NN) algorithm, analyzing historical data patterns (nearest neighbors) to generate momentum predictions.6 The primary goal is to offer a more adaptive momentum reading compared to traditional oscillators. Signals generated include trend direction and strength indications, identification of market consolidation phases (when prediction lines flatten near the mid-level), and potential overbought/oversold conditions. The user's strategy likely utilizes the crossover signal between the MLMI prediction line and its WMA.6 Key parameters include 'Prediction Data (k)', controlling the number of neighbors used (affecting sensitivity vs. stability), and 'Trend length', defining the momentum calculation window (affecting responsiveness vs. smoothness).6 The indicator's author notes potential ineffectiveness on higher timeframes (daily+) due to data limitations, reinforcing its suitability for the strategy's specified 5-minute timeframe.6

Nadaraya-Watson Rational Quadratic Kernel Regression (NW-RQK): This indicator implements a non-parametric kernel regression technique to estimate a best-fit curve for price data without assuming an underlying distribution.7 It utilizes a Rational Quadratic Kernel, which assigns higher weights to more recent data points compared to a uniform kernel (like a Simple Moving Average), allowing for quicker reaction to price changes.7 The Rational Quadratic Kernel, defined by the formula K(x,x′)=(1+∣∣x−x′∣∣2/(2αh2))−α, offers flexibility through the 'alpha' (α) hyperparameter, controlling curve smoothness, and the bandwidth 'h'.7 A critical feature is its non-repainting nature, meaning indicator values for a closed bar remain fixed, which is essential for reliable backtesting and live execution.7 Signals are derived from the estimated curve's behavior, potentially indicated by color changes based on the rate of change (slope) or crossovers. An explicit alert stream providing bullish (1) or bearish (-1) signals is also mentioned.7 The user's strategy employs the curve turning bullish or bearish as its signal. Adjustable parameters include 'Bandwidth' (lookback window size) and the 'Relative Weighting Parameter (Alpha)', which significantly impacts the curve's fit and smoothness.7

Fair Value Gap (FVG) - LuxAlgo: The LuxAlgo FVG indicator identifies areas of price imbalance between buyers and sellers, visualized as gaps on the chart.8 It employs a specific three-bar pattern logic:

A Bullish FVG is detected if the current low is above the high two bars prior (low > high(t-2)), the previous close is also above that high (close(t-1) > high(t-2)), and the relative gap height meets a threshold.8

A Bearish FVG is detected if the current high is below the low two bars prior (high < low(t-2)), the previous close is also below that low (close(t-1) < low(t-2)), and the relative gap height meets a negative threshold.8 The strategy requires an FVG mitigation signal on the 1-minute timeframe. Mitigation occurs when price returns to at least partially fill a previously identified FVG.8 For a bullish FVG, the mitigation level is the lower boundary; for a bearish FVG, it's the upper boundary. Once mitigated, the imbalance is considered resolved, potentially leading to a price reversal.8 Parameters include 'Threshold %' for filtering gap size, 'Auto Threshold' for dynamic filtering based on volatility, and the ability to specify the detection timeframe independently of the chart timeframe 8, although the user explicitly uses the 1-minute chart for FVG detection.

B. Core Logic Assessment

Mechanism: The strategy's entry logic is sequential and requires confirmation from all three indicators across two timeframes. A trade is initiated only after the following sequence completes:

One of the 5-minute indicators (MLMI or NW-RQK) generates an initial signal (e.g., MLMI bullish crossover or NW-RQK turns bullish).

A Fair Value Gap mitigation occurs on the 1-minute timeframe (confirming the direction of the initial 5-min signal).

The other 5-minute indicator provides a confirming signal (e.g., if MLMI signaled first, NW-RQK must now turn bullish). The order of steps 2 and 3 can vary, leading to four distinct valid entry sequences: MLMI → NW-RQK → FVG, MLMI → FVG → NW-RQK, NW-RQK → FVG → MLMI, and NW-RQK → MLMI → FVG.

Rationale: The strategy's architect posits that the flexibility in the sequence of the second and third signals (FVG and the second 5-min indicator) is a key source of robustness. The claim is that this lack of rigid hierarchy allows the signal sequence to adapt naturally to prevailing market dynamics. The overall goal is to capture a significant portion of a trend after its initial phase, using the multiple confirmations to filter noise and improve entry precision.

C. Initial Robustness Evaluation

Strengths:

Multi-Factor Confirmation: The requirement for agreement between momentum (MLMI), price fitting (NW-RQK), and price imbalance/liquidity (FVG) provides a potentially strong filter against random noise.

Multi-Timeframe Analysis: Integrating signals from both a 5-minute context timeframe and a 1-minute execution timeframe is a common technique to improve trade location and confirmation.9 The higher timeframe identifies the potential trend, while the lower timeframe refines the entry point.10

Potential Adaptability (Claimed): If different market conditions naturally cause the FVG and the second 5-minute indicator signals to appear in varying orders relative to the initial signal, the flexible sequencing could offer some degree of passive adaptation.

Potential Weaknesses & Considerations:

Complexity: Managing three distinct indicators, each with its own parameters and behavior, coupled with the intricate sequential triggering logic, introduces significant complexity. This increases the potential for implementation errors, optimization difficulties, and unexpected interactions. The strategy's architect suggests that traditional backtesting of the strategy as a monolithic entity might be misleading due to its dynamic nature. Instead, they propose evaluating performance by backtesting within distinct market regimes separately or employing techniques like Robust Principal Component Analysis (RPCA) to understand signal sequencing and optimize an adaptive agent based on its evolving expertise, thereby focusing on the agent's overall performance rather than a static strategy.

yes - to reduce complexity we dont need to care for those complex sequence of trading indicators - w eshould and must leverage this attitude by increasing the awareness of the usage of market regime and to investigate which seqjuence caused by which market regime ankd than perform a deep pca to understand the most likely correlation to understand the data reduce complexity and increase awareness of sxuccessful trades and replicating those trades with a far better understanding of how market dynamics would work in those situations

Indicator Lag: All technical indicators, including those incorporating machine learning, possess inherent lag relative to instantaneous price movements. Requiring a sequence of three signals, particularly across different timeframes, could compound this lag. While the strategy aims to enter after trend initiation, excessive lag might lead to entries too late in the move, reducing profit potential. The architect acknowledges this inherent lag but views it as a deliberate feature, aiming to capture trends approximately 10-15% after their initial inception. This approach, they suggest, allows for tighter stop-loss placements and greater overall efficiency, as the strategy is not designed to catch the very beginning of a move.

we capture the trend after its initiations that would raise the abilityto believe in a trend thanks to the lagging indicators - this strategy leverage those lagging issues to a daramtoclly positive directions

Parameter Sensitivity: The performance of MLMI depends on the choice of 'k' and 'trend length' 6, while NW-RQK relies on 'bandwidth' and 'alpha'.7 The assertion that the strategy is robust "without configuring the params" appears highly optimistic. Optimal performance is likely contingent on careful tuning of these parameters for the specific asset and market conditions traded, potentially contradicting the claim of parameter independence. Research often shows dynamic parameter optimization significantly outperforms fixed parameters.13 The architect clarifies that the claim of robustness without constant parameter configuration stems from manual trading experience, where real-time parameter adjustment based on regime understanding is impractical. While acknowledging that parameters could be tuned for "hyper results," the baseline strategy is intended to be functional without such continuous optimization.

no need to optimise the strategy for superior results - default params will work best across all sequences - the reason why we wont like to optimize theparams of each indicator is thatbecause we wouldlike to to casause for robustness over all periods -and if everything works properly - why change? especially when the markets improve and more capital enter to the markets we would need more stable solution and the stable solution is lifelong indicator stability will promise us superior results especkially when integrating with market regime engine and RL agent that will learn to capture the trades

Correlation Risk: The effectiveness of multi-factor confirmation depends on the relative independence of the signals. In strongly trending markets, momentum indicators (like MLMI) and trend-following regression models (like NW-RQK) might become highly correlated, providing redundant information rather than independent confirmation. However, the architect notes that even in strongly trending markets where the 5-minute indicators might show correlation, the FVG mitigation on the 1-minute timeframe serves as a crucial differentiating factor, providing an opportunity for entry during a retracement to an imbalance zone, provided all signals align.

without mitigation we have no level of confidence that we can trust perfectly

Overfitting Risk: The strategy's complexity and flexibility, especially if parameters are tuned during development, create a significant risk of overfitting to historical data.14 A strategy that looks excellent in backtests might fail in live trading if it has inadvertently learned noise rather than a genuine market edge. To address this, the architect emphasizes several points: the core strategy is intended to be robust without forcing parameter configurations, which can lead to overfitting. Trading multiple assets is proposed as a way to diversify and reduce curve-fitting to a single instrument. Furthermore, the strict rule that all three signals must align for a valid entry acts as a strong filter, focusing on higher-quality trades and inherently avoiding trades in ranging or consolidating markets where the strategy is not expected to perform, a limitation deemed acceptable.

D. Implicit Assumptions & Potential Conflicts

The strategy's design rests on certain assumptions about how its components interact, which may not always hold true. The core logic assumes that the relationship between the signals generated by MLMI, NW-RQK, and FVG mitigation remains sufficiently stable across different market conditions for the flexible sequencing to be inherently adaptive. However, market regimes are characterized by distinct patterns of volatility, correlation, and price behavior.4 A shift in regime can fundamentally alter how these indicators behave and relate to each other. For instance, during periods of high, directionless volatility ("choppy" markets), 1-minute FVGs might form frequently and appear somewhat randomly relative to the slower-developing trends captured by the 5-minute MLMI and NW-RQK indicators.6 In such a regime, an FVG mitigation signal might occur but lack a meaningful connection to the underlying directional bias the 5-minute indicators are attempting to identify. Therefore, the simple flexibility of allowing any valid sequence does not guarantee adaptation; the relevance of the sequence itself within the current market context might degrade. The strategy implicitly assumes sequence flexibility equals contextual relevance, but this link can break down if the underlying signal dynamics change drastically between regimes. An explicit mechanism to validate the meaningfulness of a detected sequence within the current regime appears necessary. The architect acknowledges that conflicts can arise but underscores that the strategy's discipline—requiring all signals to align for a "concrete trading idea"—means no trade is taken under ambiguity. The aim is not to capture every potential move but to filter for high-conviction setups.

Furthermore, a potential conflict exists between the nature of the indicators and the strategy's stated goal of capturing trends after initiation. MLMI and NW-RQK are primarily trend and momentum-following tools.6 FVGs, particularly their mitigation, often represent areas of temporary imbalance or occur during pullbacks within a larger trend.8 While entering after an initial 5-minute signal, followed by an FVG mitigation, and then confirmed by the second 5-minute signal (A → B → C sequence) might provide a well-timed entry on a pullback, the alternative sequence (A → C → B) presents a different scenario. Here, the entry occurs only when an FVG mitigation (B) happens after both 5-minute indicators (A and C) have already signaled agreement on the trend. This FVG mitigation might occur significantly later in the trend's development, potentially closer to exhaustion. Thus, the flexible sequencing intended for robustness could paradoxically lead to substantially different entry timings relative to the trend's lifecycle, with some sequences potentially resulting in entries much later than optimal for capturing a significant part of the move. Regarding the concern of an FVG mitigation occurring too soon after the initial 5-minute signal, the architect believes that if the FVG mitigation is followed by a swift confirmation from the other 5-minute indicator (e.g., MLMI aligning shortly after mitigation, or the NW-RQK curve turning bullish/bearish right after), this sequence resolves the potential conflict and validates the entry.

this is a direct relation between the vary of sequence and the direct relation of the understanding of what market regime do we have right now to map out different and irrelevant entry signals - and on consolidating markets (depends on the range the market consolidated between) we ,ight be able to execute great trades - again this is all about reinforcement leaning that will map out everything and will learn by proper training

WE INCREASE THE TIMEFRAMES TO 30 MIN THAT WILL SPOT OVER THE MLMI AND QUAD REGRESSION AND TO 5 MIN THAT WILL SPOT ON THE FVG

אני רוצה שנעמיק את המחקר הנוכחי - עם יכולות יותר אפליקטיביות - כמובן שהשפה תיהיה עברית - אני רוצה שתשסכל על כמה שיותר מאמרים והוכחותמדעיות שיכולות להוביל אותנו להבנה ברורה יותר של איך המערכת צריכה להיראות ותתייחס לכל מה שרשמתי פה