# שילוב סופרפוזיציה בהשראת קוונטים במודלים של מסחר אלגוריתמי: ניתוח והערכה

## 1. מבוא: הפרדיגמה הקוונטית-השראתית במסחר אלגוריתמי

### 1.1 המרדף אחר מודלי מסחר מתקדמים

התחום הדינמי של מסחר אלגוריתמי מאופיין במרדף מתמיד אחר אסטרטגיות מתקדמות המסוגלות לספק יתרון תחרותי, במיוחד בשווקים הפיננסיים המורכבים והתנודתיים של ימינו. מסחר אלגוריתמי, במהותו, שואף לנצל חוסר יעילות בשוק באמצעות קבלת החלטות מבוססת נתונים ובמהירות גבוהה. עם זאת, מודלים קלאסיים נתקלים לעיתים קרובות במגבלות, מה שמניע חיפוש אחר פרדיגמות חדשות. למידת חיזוק (Reinforcement Learning, RL) הופיעה כגישה מבטיחה, אך גם היא מתמודדת עם אתגרים משמעותיים, כגון חוסר יציבות של מדיניות (policy instability) ובקבוקי צוואר בתהליך הדגימה (sampling bottlenecks). אתגרים אלו מדגישים את הצורך בחדשנות מתמשכת ובחקירת גישות היכולות להתמודד טוב יותר עם אי-ודאות ודינמיות.

### 1.2 הגדרת סופרפוזיציה בהשראת קוונטים עבור מערכות חישוב קלאסיות

אלגוריתמים בהשראת קוונטים (Quantum-Inspired Algorithms, QIAs) מייצגים קטגוריה של טכניקות חישוב קלאסיות השואלות מושגים מתורת המחשוב הקוונטי במטרה לפתור בעיות מסוימות ביעילות רבה יותר מאשר שיטות קלאסיות מסורתיות. אלגוריתמים אלו מדמים עקרונות קוונטיים, כגון סופרפוזיציה ושזירה, באמצעות חומרה קלאסית, לרוב על ידי שימוש במודלים מתמטיים או בדגימה הסתברותית. בניגוד לאלגוריתמים קוונטיים אמיתיים, הדורשים חומרה קוונטית ייעודית, אלגוריתמים בהשראת קוונטים פועלים על מערכות קלאסיות סטנדרטיות, כגון מעבדים מרכזיים (CPUs) או מעבדים גרפיים (GPUs), עובדה ההופכת אותם לנגישים יותר ליישום בטווח הזמן המיידי. הדמיה קלאסית של סופרפוזיציה, למשל, יכולה לאפשר חקירה מהירה יותר של מרחבי פתרונות בבעיות אופטימיזציה מורכבות.

הנגישות של אלגוריתמים בהשראת קוונטים מהווה גשר קריטי בין הפוטנציאל התיאורטי העצום של מחשוב קוונטי לבין היישומים המעשיים האפשריים כיום. בעוד שמחשוב קוונטי אמיתי נמצא עדיין בשלבי פיתוח ומחקר אינטנסיביים , גישות QI מאפשרות לחוקרים ולמהנדסים בתעשייה, במיוחד בתחומים עתירי חישוב כמו מימון , להתחיל לחקור ולנצל עקרונות קוונטיים כבר עתה. היכולת ליישם עקרונות אלו באמצעות כלים קלאסיים זמינים אינה רק עניין של יעילות חישובית פוטנציאלית, אלא גם של דמוקרטיזציה של גישה לפרדיגמות חשיבה חדשות, מה שמאפשר התקדמות מעשית בתחום המימון גם בהיעדר מחשבים קוונטיים בקנה מידה מלא.

### 1.3 מטרת הדוח

דוח זה נועד לבצע ניתוח והערכה מעמיקים של אסטרטגיות מסחר אלגוריתמי מבוססות למידת חיזוק, אשר ממנפות את הרעיון של סופרפוזיציה בהשראת קוונטים. הדוח יבחן ארבע גישות ספציפיות: ביצוע מדיניות באנסמבל, בחירת פעולה הסתברותית, מעקב אחר השערות מצב מרובות, והתחייבות מאוחרת דמוית-QIDDM. עבור כל גישה, יוערכו איכותה והתאמתה למסחר אלגוריתמי, ובסופו של דבר, תיקבע הגישה המתאימה ביותר ליישום מעשי, בהתבסס על ניתוח חומר המחקר הזמין.

## 2. מסגרת רעיונית: הדמיית סופרפוזיציה במודלי מסחר

### 2.1 חיקוי קלאסי של סופרפוזיציה קוונטית

המושג "סופרפוזיציה" במכניקת הקוונטים מתאר את יכולתה של מערכת קוונטית, כגון קיוביט, להתקיים במספר מצבים בו-זמנית. מחשבים קלאסיים, הפועלים על ביטים שיכולים להיות רק 0 או 1, אינם יכולים לשכפל סופרפוזיציה אמיתית זו. עם זאת, אלגוריתמים בהשראת קוונטים (QIAs) מנסים לחקות היבטים מסוימים של סופרפוזיציה באמצעות טכניקות קלאסיות. הדבר נעשה לעיתים קרובות על ידי הדמיית מצבים הסתברותיים או ביצוע חישובים מקבילים. לדוגמה, מערכת קלאסית יכולה לייצג מספר פתרונות אפשריים בו-זמנית, לשמור על התפלגויות הסתברות על פני קבוצת פעולות אפשריות, או להפעיל מספר מודלים שונים במקביל, כאשר כל אחד מהם מייצג "מצב" או "השקפה" שונה.

### 2.2 העיקרון הכללי: ייצוג מצבים או פעולות פוטנציאליים מרובים במקביל לפני קבלת התחייבות

העיקרון המנחה מאחורי הדמיית סופרפוזיציה במודלים קלאסיים הוא היכולת לייצג ולעבד מספר מצבים או פעולות פוטנציאליים במקביל, לפני שהמערכת "מתחייבת" למצב או פעולה בודדים. הרעיון המרכזי הוא לאפשר למערכת "לשקול" מספר אפשרויות או "להיות" במספר מצבים פוטנציאליים בו-זמנית (באופן מדומה). גישה זו מאפשרת גמישות רבה יותר, חקירה יסודית יותר של מרחב הפתרונות, ובסופו של דבר, קבלת החלטות מושכלת יותר, במיוחד בתנאים של אי-ודאות גבוהה המאפיינים שווקים פיננסיים. מאפיין מרכזי של גישה זו הוא דחיית ההתחייבות לפעולה ספציפית עד לרגע שבו נאסף מספיק מידע כדי "לקרוס" את הסופרפוזיציה המדומה לפעולה מוגדרת אחת.

הדמיית סופרפוזיציה במערכות קלאסיות אינה רק טכניקה חישובית; היא משקפת שינוי פרדיגמטי באופן שבו ניגשים לפתרון בעיות, במיוחד בתחומים דינמיים כמו מסחר. במקום לחפש נתיב אופטימלי יחיד באופן סדרתי, המערכת חוקרת מרחב של נתיבים פוטנציאליים במקביל. היתרון המרכזי של הדמיית סופרפוזיציה אינו נעוץ בהכרח בהאצת חישובים (שעשויה להיות מוגבלת במערכות קלאסיות בהשוואה לפוטנציאל הקוונטי), אלא ביכולת לשמור על *גמישות אסטרטגית*. במקום להתחייב מוקדם מדי לפעולה בודדת או לאסטרטגיה יחידה, המערכת יכולה "להחזיק" מספר אסטרטגיות או פעולות כהשערות פעילות. יכולת זו מאפשרת למערכת להסתגל טוב יותר למידע חדש המתקבל או לשינויים פתאומיים בתנאי השוק. זהו מעבר מחשיבה דטרמיניסטית, שבה כל שלב מוביל באופן חד-משמעי לשלב הבא, לחשיבה הסתברותית ומקבילה יותר, שבה מגוון אפשרויות נשקלות בו-זמנית עד לנקודת החלטה קריטית.

## 3. גישה 1: ביצוע מדיניות באנסמבל – רתימת אינטליגנציה קולקטיבית

### 3.1 מנגנון: שילוב של מספר סוכני RL / מדיניויות

שיטות אנסמבל (Ensemble methods) בלמידת מכונה הן טכניקות המשלבות תחזיות או החלטות של מספר מודלים (הנקראים לעיתים "לומדים חלשים" או "סוכנים מרכיבים") במטרה להשיג ביצועים כוללים טובים יותר מאשר כל אחד מהמודלים המרכיבים בנפרד. בתחום למידת החיזוק, שיטות אנסמבל יכולות לשפר את ביצועי המדיניות הכוללת על ידי שילוב של פעולות נבחרות או הסתברויות פעולה המתקבלות מאלגוריתמי RL שונים המרכיבים את האנסמבל. האנסמבל יכול להיות מורכב מסוכנים שאומנו באמצעות אלגוריתמים שונים (לדוגמה, שילוב של Proximal Policy Optimization (PPO), Advantage Actor-Critic (A2C), ו-Deep Deterministic Policy Gradient (DDPG), כפי שמוצע במחקר אחד ), או מסוכנים שאומנו על תתי-קבוצות שונות של נתונים, או עם היפר-פרמטרים שונים, כל זאת במטרה לעודד גיוון בין חברי האנסמבל.

### 3.2 אנלוגיית סופרפוזיציה: האנסמבל כ"סופרפוזיציה" של לוגיקות מסחר מגוונות

ניתן לראות את גישת האנסמבל כאנלוגיה קלאסית לרעיון הסופרפוזיציה הקוונטית. כל סוכן או מדיניות בתוך האנסמבל מייצג "מצב", "השקפה" או "לוגיקת מסחר" פוטנציאלית שונה על השוק. בדומה לאופן שבו מערכת קוונטית בסופרפוזיציה מכילה מגוון מצבי בסיס אפשריים, האנסמבל "מחזיק" בו-זמנית מגוון של אסטרטגיות או גישות לקבלת החלטות. ההחלטה הסופית של האנסמבל, המתקבלת באמצעות מנגנון איחוד כלשהו, היא למעשה "קריסה" של סופרפוזיציית הלוגיקות הללו לפעולה קונקרטית אחת. הגיוון בין הסוכנים המרכיבים את האנסמבל הוא קריטי להצלחת הגישה, שכן הוא מאפשר כיסוי רחב יותר של מרחב האסטרטגיות האפשריות ומפחית את הסיכון להסתמכות על נקודת מבט יחידה שעלולה להיות שגויה או מוגבלת.

### 3.3 היבטים מרכזיים

הצלחתה של גישת האנסמבל תלויה במספר היבטים מרכזיים:

* **גיוון סוכנים (Agent Diversity):** יצירת גיוון בין הסוכנים באנסמבל היא חיונית. ניתן להשיג זאת על ידי אימון סוכנים שונים באופן עצמאי , שימוש באלגוריתמי RL שונים בעלי מאפיינים הטרוגניים , או יישום שיטות ספציפיות המעודדות גיוון באופן אקטיבי. מחקר בנושא MARL-Focal, למשל, מציג גישה לאופטימיזציה של גיוון מוקדי במערכות מרובות סוכנים המבוססות על מודלי שפה גדולים (LLM), גישה שעשויה להיות בעלת פוטנציאל יישום רחב יותר גם בהקשרים של סוכני מסחר. החשיבות של גיוון נובעת מכך שאנסמבלים המורכבים ממודלים מגוונים נוטים להיות חזקים יותר, מכיוון שטעויות או הטיות של מודל אחד יכולות להתקזז על ידי החוזקות של מודלים אחרים.
* **פרוטוקולי אימון וולידציה:** פרוטוקולים סטנדרטיים כוללים אימון של כל סוכן באופן עצמאי על חלונות אימון מתגלגלים של נתונים היסטוריים (לדוגמה, חלון של 30 יום או חלון של n חודשים ). לאחר שלב האימון, מתבצעת ולידציה על חלונות נתונים נפרדים ועוקבים (למשל, 5 ימים או 3 חודשים ). שלב הוולידציה משמש לבחירת הסוכנים בעלי הביצועים הטובים ביותר או לחישוב משקלים עבור כל סוכן באנסמבל.
* **חישוב משקלים (Weight Calculation):** קיימות מספר גישות לקביעת המשקל של כל סוכן בהחלטת האנסמבל. גישה נפוצה אחת היא פסילת סוכנים שהפגינו ביצועים נמוכים מאוד (למשל, יחס שארפ נמוך מאוד) במהלך תקופת הוולידציה. עבור הסוכנים הנותרים, ניתן לחשב משקלים באמצעות פונקציית softmax המיושמת על מדדי הביצועים שלהם (כגון יחס שארפ) מתקופת הוולידציה. גישה זו מעניקה משקל גבוה יותר לסוכנים שהוכיחו ביצועים טובים יותר מותאמי סיכון. גישה אחרת, פשוטה יותר, היא בחירת הסוכן היחיד שהשיג את יחס השארפ הגבוה ביותר בתקופת הוולידציה, והוא זה שיבצע את המסחר בתקופה העוקבת (למשל, הרבעון הבא).
* **איחוד החלטות (Decision Fusion):** לאחר שכל סוכן באנסמבל מייצר המלצת פעולה או התפלגות הסתברויות על פני פעולות, יש לאחד החלטות אלו להחלטה סופית אחת. מספר מנגנוני איחוד נפוצים כוללים:
  + **ממוצע משוקלל של הסתברויות פעולה:** האנסמבל פועל על בסיס ממוצע משוקלל של הסתברויות הפעולה שסופקו על ידי כל אחד מהסוכנים המרכיבים אותו. זהו סוג של מיזוג "רך" של החלטות, המביא בחשבון את רמת הביטחון של כל סוכן בכל פעולה.
  + **הצבעת רוב (Majority Voting):** הפעולה שנבחרה על ידי רוב הסוכנים באנסמבל נבחרת כפעולת האנסמבל הסופית (עבור מסחר בקריפטו) (עבור EMAX). זהו מנגנון מיזוג "קשה".
  + **בחירת הסוכן הטוב ביותר:** כפי שתואר קודם, הסוכן בעל הביצועים הטובים ביותר בתקופת הוולידציה נבחר לבצע את המסחר.
  + **איחוד מבוסס מסווגים/סינון שונות:** מחקרים עדכניים מציעים לשלב סוכני RL עם מסווגים קלאסיים, תוך שימוש במנגנוני סינון המבוססים על שונות ציוני הביטחון של הסוכנים השונים, ובחירת פעולה אדפטיבית המביאה בחשבון את רמת ההסכמה בין הסוכנים.
  + **איחוד היררכי:** ארכיטקטורות מתקדמות יותר יכולות לכלול סוכן-על (meta-agent) הלומד כיצד לשקלל בצורה אופטימלית את החלטותיהם של סוכני המשנה באנסמבל, כפי שהוצע בארכיטקטורת אנסמבל DQN המשלבת CNN ו-LSTM.

### 3.4 הערכה למסחר

* **חוזקות:**
  + **חוסן משופר (Enhanced Robustness) ויציבות:** אחת החוזקות המרכזיות של שיטות אנסמבל היא יכולתן להפחית את השפעת הכישלונות, ההטיות והרגישות של מודלים בודדים. על ידי שילוב מספר נקודות מבט, האנסמבל יכול להסתגל טוב יותר למצבי שוק משתנים ולהפגין יציבות רבה יותר בביצועיו. מחקרים מראים שאנסמבלים יכולים להפחית משמעותית את שונות התשואות בהשוואה לסוכנים מרכיבים בודדים, ובכך למתן את חוסר יציבות המדיניות האופייני ל-RL.
  + **הפחתת התאמת יתר (Reduced Overfitting):** טכניקות אנסמבל נפוצות, כגון Bagging (Bootstrap Aggregating) ו-Boosting, מסייעות במניעת התאמת יתר של המודל לנתוני האימון, ובכך משפרות את יכולת ההכללה שלו לנתונים חדשים ובלתי נראים.
  + **שיפור תשואות מותאמות סיכון:** מספר מחקרים מדווחים על שיפורים משמעותיים במדדי ביצועים מותאמי סיכון, כגון עלייה ביחס שארפ והפחתה ב-Maximum Drawdown (משיכה מקסימלית). לדוגמה, מחקרים מסוימים הראו הפחתה של עד 4.17% ב-Maximum Drawdown ושיפור של עד 0.21 ביחס שארפ. מחקר אחר דיווח על יחס שארפ של 1.30 עבור אסטרטגיית אנסמבל, בהשוואה לטווח של 0.87-1.12 עבור הסוכנים הבודדים שהרכיבו אותו.
  + **טיפול במורכבות וברעש:** אנסמבלים מסוגלים ללכוד דפוסים מורכבים יותר בנתונים ולפצות על חולשות אינהרנטיות של מודלים בודדים, מה שהופך אותם ליעילים יותר בסביבות רועשות ומורכבות כמו שווקים פיננסיים.
* **חולשות:**
  + **תקורה חישובית (Computational Overhead):** אימון, ולידציה ותחזוקה של מספר מודלים במקביל דורשים משאבי חישוב וזמן רבים יותר בהשוואה למודל יחיד.
  + **מורכבות כוונון (Tuning Complexity):** תהליך הכוונון של היפר-פרמטרים עבור כל אחד מהמודלים באנסמבל, כמו גם עבור מנגנון האנסמבל עצמו (למשל, שיטת השקלול או איחוד ההחלטות), יכול להיות מורכב ומאתגר.
  + **תיאום שגוי פוטנציאלי (Potential Miscoordination):** במיוחד במערכות מרובות סוכנים (MARL), כאשר מספר סוכנים לומדים ופועלים במקביל, האימון המקביל של מדיניויות יכול להפוך את סביבת הלמידה ללא-נייחת (non-stationary) עבור כל סוכן, מה שעלול להוביל להערכות ערך לא יציבות, גרדיאנטים בעלי שונות גבוהה, ובסופו של דבר לתיאום שגוי בין הסוכנים. גישת EMAX (Ensemble Value Functions for Multi-Agent Exploration) מנסה לטפל בבעיה זו על ידי שימוש בממוצע הערכות ערך על פני האנסמבל ובהצבעת רוב בשלב ההערכה.
  + **רגישות לסף השונות (Sensitivity to Variance Threshold):** במנגנוני איחוד החלטות המשתמשים בסינון מבוסס שונות (כמו זה המוצע ב-), ביצועי האנסמבל יכולים להיות רגישים לבחירת סף השונות (τ) המשמש להערכת אמינות ציוני הביטחון. הדבר מדגיש את החשיבות הפוטנציאלית של התאמה דינמית של סף זה.

### 3.5 תובנות ביצועים ממחקר והצלחה פוטנציאלית

הצלחתן של שיטות אנסמבל במסחר אלגוריתמי, ובפרט בהקשר של הדמיית סופרפוזיציה, תלויה באופן קריטי באיזון העדין בין *גיוון* המדיניויות המרכיבות את האנסמבל לבין *קוהרנטיות* ההחלטה הסופית המאוחדת. גיוון יתר, ללא מנגנון איחוד החלטות חזק ומתוחכם, עלול להוביל לרעש, חוסר החלטיות, ואף לביצועים נמוכים יותר (בדומה לתופעת דה-קוהרנטיות במערכות קוונטיות, שבה האינטראקציה עם הסביבה הורסת את הסופרפוזיציה העדינה). מנגד, חוסר גיוון מספק (למשל, אם כל הסוכנים באנסמבל מתכנסים לאותה אסטרטגיה או לוגיקה דומה מאוד) מבטל למעשה את יתרון הסופרפוזיציה המדומה, והאנסמבל יתנהג כמו סוכן יחיד.

האתגר המרכזי, אם כן, אינו רק ביצירת קבוצה של סוכנים מגוונים, אלא בעיצוב מנגנון איחוד החלטות (או "מדידה" קלאסית, בהשאלה מהקוונטים) שממקסם את ה"סיגנל" הקולקטיבי הנובע מהחוכמה המשותפת של הסוכנים, תוך מזעור ה"רעש" הנובע מחילוקי דעות, טעויות אינדיבידואליות של סוכנים, או תיאום שגוי. שיטות כמו חישוב משקלים מבוסס יחס שארפ או בחירת הסוכן בעל הביצועים הטובים ביותר בוולידציה הן ניסיונות ראשוניים למטב את תהליך ה"קריסה" הזה של הסופרפוזיציה המדומה לפעולה אחת. גישות מתקדמות יותר, כגון EMAX המשתמש בהצבעת רוב כדי להפחית תיאום שגוי , או מנגנוני איחוד מבוססי מסווגים עם סינון שונות של ציוני ביטחון , מנסות לטפל בבעיית הקוהרנטיות והאמינות של ההחלטה הסופית באופן מתוחכם יותר.

הטבלה הבאה מסכמת מדדי ביצועים נבחרים עבור שיטות אנסמבל במסחר אלגוריתמי, כפי שדווחו במחקרים שנסקרו:

| שיטת אנסמבל / סוכן | מדד מפתח | ערך מדווח | השוואה (מקור) | קטע מחקר |
| --- | --- | --- | --- | --- |
| אנסמבל (ממוצע הסתברויות PPO, SAC, DDPG) | Max Drawdown (הפחתה) | 4.17% | ביחס לסוכנים בודדים |  |
| אנסמבל (ממוצע הסתברויות PPO, SAC, DDPG) | יחס שארפ (שיפור) | 0.21 | ביחס לסוכנים בודדים |  |
| אנסמבל (בחירת הטוב ביותר: PPO, A2C, DDPG) | יחס שארפ | 1.30 | PPO: 1.10, A2C: 1.12, DDPG: 0.87 |  |
| אנסמבל (בחירת הטוב ביותר: PPO, A2C, DDPG) | Max Drawdown | -9.7% | PPO: -23.7%, A2C: -10.2%, DDPG: -14.8% |  |
| אנסמבל (ממוצע הסתברויות PPO, SAC, DDPG) | סטיית תקן תשואות | ~50% נמוכה יותר | ביחס לסוכנים בודדים |  |
| אנסמבל RL עם מסווגים (שיטות שונות) | ביצועים כלליים | עליונות עקבית | ביחס למודלי בסיס (SR, MDD, Calmar) |  |

טבלה זו מרכזת ראיות אמפיריות התומכות ביעילותה של גישת האנסמבל, ומדגישה את היתרונות הכמותיים הפוטנציאליים במונחי תשואה מותאמת סיכון וחוסן.

## 4. גישה 2: בחירת פעולה הסתברותית – ניווט באי-ודאות באמצעות הסתברויות

### 4.1 מנגנון: מדיניויות המייצרות התפלגויות הסתברות על פני פעולות מסחר דיסקרטיות

בגישת בחירת פעולה הסתברותית, במקום שהמדיניות של סוכן הלמידה (agent's policy) תקבע באופן דטרמיניסטי פעולה אחת ויחידה שיש לבצע במצב נתון, היא פולטת התפלגות הסתברות על פני כל הפעולות האפשריות במרחב הפעולה. לדוגמה, במסחר, מרחב הפעולה יכול להיות דיסקרטי ולכלול פעולות כגון 'קנה' (Long), 'מכור' (Short), או 'החזק' (Flat/Hold). המדיניות, המיוצגת לרוב על ידי רשת נוירונים, מקבלת כקלט את מצב השוק הנוכחי (למשל, נתוני מחירים היסטוריים, אינדיקטורים טכניים) ומוציאה וקטור של הסתברויות, כאשר כל רכיב בווקטור מייצג את ההסתברות לבחירת הפעולה המתאימה. הפעולה הסופית שיבצע הסוכן נדגמת מהתפלגות זו. גישה זו מאפשרת באופן טבעי איזון בין חקירה (exploration) של פעולות חדשות או פחות מוכרות, לבין ניצול (exploitation) של ידע קיים לגבי פעולות שהוכחו כיעילות בעבר.

יתרה מכך, ייצוג בעיות קבלת החלטות סדרתיות, כגון אלו הניצבות בפני סוכן מסחר, כבעיות של הסקה הסתברותית (probabilistic inference) יכול להיות בעל ערך רב. מסגרת ההסקה ההסתברותית מציעה מגוון רחב של כלים מתמטיים רבי עוצמה להסקת הדינמיקה הסטוכסטית של הסביבה (כלומר, השוק) ובמקביל מציעה פרשנות הסתברותית לתהליך אופטימיזציית המדיניות עצמו.

### 4.2 אנלוגיית סופרפוזיציה: התפלגויות הסתברות מייצגות באופן אינהרנטי סופרפוזיציה של פעולות

התפלגות ההסתברות על פני הפעולות האפשריות מהווה אנלוגיה קלאסית לסופרפוזיציה קוונטית. כל פעולה אפשרית (למשל, קנה, מכור, החזק) "קיימת" במערכת עם "משרעת" (amplitude) מסוימת, המיוצגת על ידי ההסתברות שלה. לפני שמתבצעת "מדידה" – כלומר, לפני שנדגמת פעולה ספציפית מההתפלגות – המערכת נמצאת למעשה ב"סופרפוזיציה" של כל הפעולות הפוטנציאליות, כאשר כל פעולה משוקללת על ידי ההסתברות שלה. תהליך הדגימה הוא המקביל הקלאסי ל"קריסת פונקציית הגל" במכניקה הקוונטית, שבו הסופרפוזיציה מתורגמת לתוצאה מוגדרת אחת.

### 4.3 יצירה והערכה

* **פלט רשתות נוירונים:** רשתות נוירונים, ובפרט כאלו המשתמשות בפונקציית הפעלה מסוג softmax בשכבת הפלט, משמשות באופן נפוץ ליצירת התפלגויות הסתברות אלו על פני מרחב פעולות דיסקרטי. הקלט לרשת הוא ייצוג של מצב השוק הנוכחי, והפלט הוא וקטור הסתברויות התואם את הפעולות האפשריות.
* **כיול תחזיות וחדות (Calibration and Sharpness):** כאשר הפעולות ההסתברותיות מבוססות על תחזיות מחיר הסתברותיות (למשל, הסתברות שהמחיר יעלה, ירד או יישאר יציב), איכות התחזיות הללו הופכת לקריטית. שני מדדים חשובים להערכת תחזיות הסתברותיות הם כיול וחדות.
  + **כיול (Calibration):** מודד עד כמה ההסתברויות החזויות תואמות את התוצאות הנצפות בפועל. לדוגמה, אם המודל חוזה הסתברות של 70% לעליית מחיר, האם המחיר אכן עולה ב-70% מהמקרים שבהם ניתנה תחזית זו? מודל מכויל היטב יציג התאמה גבוהה בין ההסתברויות החזויות לתדירויות הנצפות.
  + **חדות (Sharpness):** מודדת עד כמה התפלגות ההסתברויות החזויה היא "מרוכזת" או "מפוזרת". התפלגות חדה יותר (כלומר, צרה יותר, עם הסתברויות גבוהות המוקצות למספר קטן של תוצאות) מעידה על ביטחון גבוה יותר בתחזית. ניתן להשתמש בטכניקות מתקדמות כגון Conformal Prediction (CP) כדי לכמת אי-ודאות בצורה סטטיסטית קפדנית וליצור מרווחי חיזוי (prediction intervals) שהם קליברים, כלומר, מכסים את התוצאה האמיתית בשכיחות הרצויה.

### 4.4 הערכה למסחר

* **חוזקות:**
  + **טיפול אינהרנטי באי-ודאות:** הגישה ההסתברותית מכירה באופן מובנה בכך שהעתיד אינו ידוע בוודאות ומספקת מסגרת טבעית להתמודדות עם אי-ודאות זו. על ידי ייצוג החלטות כהתפלגויות, המודל יכול לשקף את רמת אי-הוודאות שלו.
  + **גמישות:** מדיניויות הסתברותיות מאפשרות למודל להסתגל לתנאי שוק משתנים על ידי התאמה דינמית של הסתברויות הפעולה בתגובה למידע חדש.
  + **איזון בין חקירה לניצול (Exploration/Exploitation):** תהליך הדגימה מהתפלגות המדיניות מאפשר באופן טבעי לחקור פעולות שאינן נראות אופטימליות באופן מיידי (כלומר, בעלות הסתברות נמוכה יותר) אך עשויות להתגלות ככאלה בטווח הארוך. זהו מרכיב חיוני בלמידת חיזוק יעילה.
  + **קבלת החלטות ללא רגש:** בדומה לכל אסטרטגיית מסחר אלגוריתמי, גם גישה זו מבוססת על לוגיקה קבועה מראש ואינה מושפעת מהטיות רגשיות אנושיות.
* **חולשות:**
  + **פרשנות של פלטים הסתברותיים:** לעיתים קרובות קשה לפרש מדוע המודל הקצה הסתברות מסוימת לפעולה ספציפית במצב נתון. חוסר שקיפות זה, המכונה לעיתים בעיית "הקופסה השחורה", עלול להקשות על בניית אמון במודל ועל איתור מקורן של טעויות (מציין חוסר שקיפות כאתגר כללי באלגוריתמים).
  + **סיכון להתאמת יתר להתפלגויות היסטוריות:** אם ההתפלגויות הנלמדות על ידי המודל משקפות באופן הדוק מדי את דפוסי העבר, הן עלולות שלא להכליל היטב לתנאי שוק עתידיים השונים באופן מהותי מאלו שנצפו במהלך האימון (מציין התאמת יתר כאחת החסרונות הפוטנציאליים של מסחר אלגוריתמי).
  + **אתגרים ב-MARL (Multi-Agent Reinforcement Learning):** ביישומים הכוללים מספר סוכנים הלומדים ופועלים במקביל (למשל, מספר אסטרטגיות המתחרות או משתפות פעולה), השימוש במדיניויות הסתברותיות יכול להוביל לחוסר יציבות, לצורך בתיאום מורכב בין הסוכנים, ולקשיים בהקצאת "אשראי" (credit assignment) – כלומר, קביעה אילו פעולות של אילו סוכנים תרמו לתוצאה הכוללת.

### 4.5 תובנות ביצועים ממחקר והשלכות פילוסופיות

השימוש במדיניות הסתברותית במסחר אלגוריתמי מבוסס RL הוא למעשה הכרה מובנית בכך שבשווקים פיננסיים, לעיתים רחוקות קיימת "פעולה נכונה" אחת ויחידה בכל רגע נתון. במקום זאת, קיים מרחב של פעולות אפשריות, שלכל אחת מהן סבירות שונה להוביל לתוצאה רצויה. ניתן לראות זאת כשיקוף של עקרון אי-הוודאות של הייזנברג ברמה פילוסופית – ככל שאנו מנסים להיות בטוחים יותר בפעולה אחת ספציפית, אנו עלולים להתעלם מאפשרויות אחרות שעשויות להיות אופטימליות יותר בטווח הארוך או בתרחישים מסוימים.

מדיניות הסתברותית אינה רק כלי לניהול אי-ודאות; היא גם מנגנון מובנה ללמידה מתמשכת והסתגלות. על ידי כך שהיא, באופן עקרוני, אף פעם לא פוסלת לחלוטין פעולות (כל עוד יש להן הסתברות שונה מאפס), היא משאירה "דלת פתוחה" לגילוי אסטרטגיות חדשות או להסתגלות לשינויי משטר בשוק שאסטרטגיה דטרמיניסטית, הנוטה להתקבע על פעולה אופטימלית יחידה, עלולה לפספס. האתגר המרכזי טמון בכיול נכון של הסתברויות אלו , כך שישקפו באופן מהימן את הסיכויים והסיכונים הכרוכים בכל פעולה, וכן באיזון הנכון בין חקירה (המאפשרת גילוי) לניצול (המאפשר רווחיות). מחקרים מראים כי ניתן לעדכן את המדיניות כך שפעולות המובילות לתגמולים מעל הממוצע הופכות לסבירות יותר לאורך זמן , ותחזיות מחיר הסתברותיות, שעליהן יכולה להתבסס המדיניות, מציעות מבט מקיף יותר על אי-ודאות מאשר הערכות נקודתיות דטרמיניסטיות.

## 5. גישה 3: מעקב אחר השערות מצב מרובות – פענוח דינמיקות שוק עמומות

### 5.1 מנגנון: שמירה והערכה במקביל של מספר השערות לגבי מצבי שוק בסיסיים או מסלולים עתידיים

גישת מעקב אחר השערות מצב מרובות (Multi-Hypothesis State Tracking, MHST), ובפרט אלגוריתם מעקב רב-השערות (Multi-Hypothesis Tracking, MHT), נחשבת לאחת משיטות המעקב המתקדמות ביותר. היא כוללת יצירה וניהול של עץ של "מסלולים" או "השערות" פוטנציאליים. בהקשר של מסחר אלגוריתמי, כל השערה כזו יכולה לייצג מצב שוק בסיסי אפשרי (למשל, "שוק שורי", "שוק דובי", "שוק תנודתי"), דינמיקה עתידית צפויה, או אפילו אסטרטגיה מתחרה שפעילה בשוק. האלגוריתם מעריך את הסבירות של כל השערה לאור נתונים נכנסים, ומבצע "גיזום" (pruning) של השערות שהופכות לבלתי סבירות. גישה זו יעילה במיוחד בשמירה על עקביות במעקב אחר מספר "אובייקטים" (מצבי שוק) בו-זמנית, ובטיפול באי-ודאויות ועמימות הנגרמות מרעש או מידע חלקי בנתוני השוק.

אחת הדרכים ליישם מעקב אחר מצבי שוק חבויים היא באמצעות מודלים סטטיסטיים כגון מודל מרקוב חבוי (Hidden Markov Model, HMM). מודלי HMM מאפשרים לזהות "משטרי שוק" (market regimes) חבויים, שאינם נצפים ישירות, על סמך התנהגות של משתנים נצפים (כגון מחירי נכסים, נפחי מסחר, או אינדיקטורים כלכליים).

### 5.2 אנלוגיית סופרפוזיציה: קיום מקביל של מספר השערות מצב שוק

ניתן לראות את תהליך המעקב אחר השערות מרובות כאנלוגיה קלאסית נוספת לסופרפוזיציה. כל השערה לגבי מצב השוק (למשל, "השוק נמצא במגמת עלייה חזקה", "השוק נכנס לתקופת דשדוש עם תנודתיות נמוכה", "אסטרטגיית X מבוססת מומנטום היא הדומיננטית כרגע") "קיימת" במערכת עם סבירות או "משקל אמונה" (belief weight) מסוים. המערכת עוקבת אחר כל ההשערות הללו בו-זמנית, וכל פיסת מידע חדשה מהשוק (למשל, פרסום נתון מאקרו-כלכלי, תנועת מחיר חריגה) משמשת לעדכון הסבירויות של כל אחת מההשערות. בדומה לסופרפוזיציה קוונטית, שבה קיימים מספר מצבים בו-זמנית עד ל"מדידה", כאן קיימות מספר השערות בו-זמנית עד שהראיות מצטברות מספיק כדי "לקרוס" את האמונה למצב שוק דומיננטי אחד (או מספר קטן של מצבים סבירים).

### 5.3 רכיבים מרכזיים

יישום יעיל של MHST כולל מספר רכיבים מרכזיים:

* **יצירת השערות (Hypothesis Generation):** זיהוי ראשוני של מצבי שוק פוטנציאליים או אסטרטגיות מתחרות. תהליך זה יכול להתבסס על ניתוח נתונים היסטוריים, שימוש באינדיקטורים טכניים, מודלים סטטיסטיים (כמו HMM), או אפילו ידע מומחה.
* **אסוציאציית נתונים (Data Association):** השלב הקריטי של קישור תצפיות שוק חדשות (למשל, תנועות מחירים עדכניות, שינויים בנפחי מסחר, פרסום חדשות רלוונטיות) להשערות הקיימות. מטרת שלב זה היא לעדכן את הסבירות או את משקל האמונה של כל השערה לאור המידע החדש.
* **אסטרטגיות גיזום (Pruning Strategies):** כדי לשמור על יעילות חישובית ולמנוע התפוצצות קומבינטורית של מספר ההשערות, יש ליישם אסטרטגיות גיזום. אסטרטגיות אלו מסירות השערות שהופכות לבלתי סבירות (כלומר, סבירותן יורדת מתחת לסף מסוים) לאור הנתונים החדשים שנצפו.

### 5.4 הערכה למסחר

* **חוזקות:**
  + **פוטנציאל לחוסן בשווקים רועשים או עמומים:** היכולת לשקול מספר הסברים אפשריים למצב השוק הנוכחי יכולה להוביל לקבלת החלטות מסחר חזקות יותר, במיוחד כאשר האותות מהשוק חלשים, רועשים או סותרים.
  + **זיהוי מוקדם של שינויי משטר:** על ידי מעקב מתמיד אחר הסבירות של משטרי שוק שונים, המערכת עשויה לזהות מעברים בין משטרים (למשל, מעבר משוק שורי לשוק דובי) בשלב מוקדם יותר מאשר מודלים המתמקדים במצב יחיד או מגיבים לשינויים רק לאחר שהם כבר התבססו. מחקרים המשלבים HMM עם סוכני RL לזיהוי משטרים מראים שיפור בביצועי המסחר.
  + **הסתגלות דינמית:** למידת חיזוק בייסיאנית (Bayesian Reinforcement Learning, BRL) מספקת מסגרת טבעית לשמירה על התפלגויות אמונה מפורשות על פני פרמטרים של מודל, כולל מצבי שוק חבויים. אמונות אלו מתעדכנות באופן רציף עם קבלת נתונים חדשים, מה שמאפשר למערכת להסתגל לשינויים. גם מסגרת ה-Dynamic Embedding Reinforcement Learning (DERL) משתמשת באנקודרים גנרטיביים ובמטא-למידה מקוונת כדי להטמיע באופן דינמי מידע שוק ולהסתגל לשינויים, כולל שינויי משטר.
* **חולשות:**
  + **דרישות חישוביות:** שמירה, עדכון והערכה של מספר רב של השערות במקביל עלולות להיות תובעניות מאוד מבחינה חישובית, במיוחד אם כל השערה דורשת מודל מורכב משלה או אם מרחב המצבים גדול.
  + **הגדרת פונקציות סבירות וכללי גיזום יעילים:** קביעת האופן המדויק שבו נתונים חדשים משפיעים על סבירות ההשערות השונות, והגדרת קריטריונים אופטימליים לגיזום השערות, הן משימות לא טריוויאליות הדורשות מומחיות וכיול קפדני.
  + **קללת המימד (Curse of Dimensionality):** ככל שמספר המאפיינים המשמשים להגדרת מצב שוק גדל (למשל, מספר האינדיקטורים הטכניים, הנתונים המאקרו-כלכליים, וכו'), מספר ההשערות הפוטנציאליות עלול לגדול באופן אקספוננציאלי, מה שהופך את הבעיה לבלתי ניתנת לניהול חישובי.

### 5.5 רלוונטיות של RL מבוסס HMM לאסטרטגיות תלויות משטר

השילוב של מודלי מרקוב חבויים (HMM) עם למידת חיזוק (RL) מציע גישה מבטיחה לפיתוח אסטרטגיות מסחר תלויות-משטר. מחקר אחד מציג מסגרת RL-HMM המאפשרת לסוכן להחליף בין אסטרטגיות למידה שונות (למשל, בין קבלת החלטות מבוססת RL לבין בחירות אקראיות או היוריסטיקה פשוטה) בהתבסס על מצבים חבויים (משטרים) הנלמדים ומתעדכנים על ידי רכיב ה-HMM. גישה זו יכולה להיות אנלוגית לסוכן מסחר המתאים את רמת האגרסיביות שלו או את סוג האסטרטגיה שהוא מפעיל בהתאם למצב השוק הנתפס (למשל, שוק תנודתי לעומת שוק מגמתי).

מחקר אחר מדגים באופן ספציפי את השילוב של HMM עם מגוון אלגוריתמי RL נפוצים (A2C, PPO, SAC, TD3, DDPG) במסגרת FinRL. במודל זה, המודעות למשטר השוק (כפי שזוהה על ידי ה-HMM) מוטמעת ישירות במרחב התצפית (observation space) של סוכן ה-RL. התוצאות מראות כי סוכני ה-RL-HMM השיגו ביצועים עדיפים על פני מודלי הבסיס (ללא מודעות למשטר) במדדים כגון יחס שארפ, תשואות שנתיות, ומדדי ביצועים אחרים מותאמי סיכון.

מעקב אחר השערות מצב מרובות, במיוחד כאשר הוא משולב עם טכניקות כמו למידת חיזוק בייסיאנית או מודלי מרקוב חבויים, מאפשר למערכת המסחר לא רק להגיב לשינויים שכבר התרחשו בשוק, אלא במידה מסוימת גם *לצפות* אותם. זאת על ידי הערכה דינמית של הסבירות למעברים בין משטרי שוק שונים. זהו מעבר מראייה ריאקטיבית, שבה המערכת מגיבה לאירועים לאחר התרחשותם, לראייה פרואקטיבית יותר של דינמיקת השוק. היכולת להחזיק "אמונות" הסתברותיות לגבי מספר משטרי שוק פוטנציאליים (שהיא למעשה הסופרפוזיציה הקלאסית של משטרים אפשריים) ולעדכן אותן באופן דינמי, מאפשרת למערכת המסחר להיות "מוכנה" למספר תרחישים עתידיים. לדוגמה, אם ההסתברות למעבר ממשטר של תנודתיות נמוכה למשטר של תנודתיות גבוהה מתחילה לעלות (גם אם המעבר טרם הושלם), המערכת יכולה להתחיל להתאים את האסטרטגיה שלה מבעוד מועד – למשל, על ידי הקטנת חשיפות, הגדלת החזקות במזומן, או מעבר לאסטרטגיות המתאימות יותר לתנודתיות גבוהה. זהו יתרון משמעותי על פני מודלים שמניחים משטר יחיד וקבוע, או שמגיבים לשינויי משטר רק לאחר שהם כבר באו לידי ביטוי מלא בשוק.

## 6. גישה 4: התחייבות מאוחרת דמוית QIDDM – אופטימיזציה של החלטות באמצעות פעולה מושהית

### 6.1 מנגנון: אלגוריתם קבלת החלטות דינמי בהשראת קוונטים (QIDDM)

אלגוריתם קבלת החלטות דינמי בהשראת קוונטים (Quantum-Inspired Dynamic Decision-Making, QIDDM) הוא מסגרת חישובית חדשנית הממנפת עקרונות ממכניקת הקוונטים – ובפרט את רעיונות הסופרפוזיציה, השזירה (entanglement, אם כי באופן מושגי יותר בהקשר הקלאסי), והקריסה (collapse) – במטרה לשפר את תהליך קבלת ההחלטות בסביבות דינמיות המאופיינות באי-ודאות גבוהה. הרעיון המרכזי מאחורי QIDDM הוא לשמור על סופרפוזיציה הסתברותית של פעולות פוטנציאליות אפשריות, ולא להתחייב לאף אחת מהן בטרם עת. ההתחייבות לפעולה ספציפית ("קריסה") מתרחשת רק כאשר נתונים קונטקסטואליים מהסביבה מספקים מספיק מידע כדי להצדיק החלטה. גישה זו מאפשרת תגובות אדפטיביות יותר לסביבה משתנה ודוחה התחייבויות מוקדמות שעלולות להתברר כשגויות.

### 6.2 אנלוגיית סופרפוזיציה: שמירה קלאסית על סופרפוזיציה הסתברותית של פעולות, עם משרעות המתעדכנות על ידי נתונים קונטקסטואליים

QIDDM מיישם את אנלוגיית הסופרפוזיציה באופן קלאסי על ידי ייצוג מרחב ההחלטות כסופרפוזיציה של N מצבי החלטה (או פעולות) אפשריים. ניתן לבטא זאת מתמטית כ: \Psi = \sum\_{i=1}^{N} \alpha\_i |s\_i\rangle כאשר |s\_i\rangle מייצג את מצב ההחלטה (הפעולה) ה-i, ו-\alpha\_i היא ה"משרעת" (amplitude) הקלאסית המשויכת למצב זה. ההסתברות לבחירת מצב החלטה ספציפי s\_i נתונה על ידי ריבוע המשרעת שלו: P(s\_i) = |\alpha\_i|^2. סכום ריבועי כל המשרעות חייב להסתכם ל-1, בדומה לנרמול של פונקציית גל קוונטית.

בהקשר של מסחר פיננסי, מצבי ההחלטה s\_i יכולים להיות, לדוגמה, {קנה, מכור, החזק}. מאפיין מרכזי של QIDDM הוא שהמשרעות \alpha\_i אינן קבועות, אלא מתעדכנות באופן דינמי באמצעות משקלים קונטקסטואליים (C) מהסביבה. לדוגמה, מדדי תנודתיות שוק, סנטימנט חדשותי, או אינדיקטורים טכניים יכולים לשמש כקלט קונטקסטואלי לעדכון המשרעות: \alpha\_i' = f(\alpha\_i, C). עדכון דינמי זה מאפשר למערכת להתאים את "נטייתה" כלפי כל פעולה אפשרית בזמן אמת.

### 6.3 "קריסת" החלטה: מופעלת על ידי עלייה במידע (הפחתת אנטרופיה)

המעבר מהסופרפוזיציה של פעולות פוטנציאליות להחלטה דטרמיניסטית אחת ("קריסה") אינו מתרחש בכל צעד זמן, אלא מופעל על ידי קריטריון המבוסס על כמות המידע הזמין. QIDDM משתמש במושג האנטרופיה של התפלגות ההסתברויות על פני הפעולות כדי לכמת את אי-הוודאות. האנטרופיה (H) מחושבת כ: H = - \sum\_{i=1}^{N} P(s\_i) \log\_2 P(s\_i) ההחלטה "קורסת" לפעולה ספציפית (לרוב, הפעולה בעלת ההסתברות הגבוהה ביותר) רק כאשר האנטרופיה H יורדת מתחת לסף (\theta) שנקבע מראש, כלומר H < \theta. סף אנטרופיה נמוך משמעו שהמערכת דורשת רמת ודאות גבוהה (כלומר, התפלגות הסתברויות "חדה" מאוד, שבה פעולה אחת דומיננטית בבירור) לפני שהיא מתחייבת לפעולה. גישה זו מבטיחה שהחלטות מתעכבות עד שנאסף מספיק מידע מהסביבה כדי להפחית את אי-הוודאות לרמה קבילה.

### 6.4 הערכה למסחר

* **חוזקות:**
  + **הפחתת התחייבויות מוקדמות בסביבות תנודתיות:** היתרון המרכזי של QIDDM הוא יכולתו לדחות קבלת החלטות עד לרגע שבו יש מספיק מידע תומך. יכולת זו מפחיתה משמעותית את הסיכון לביצוע טעויות יקרות, במיוחד בשווקים פיננסיים הידועים בתנודתיותם וב"רעש" הרב המאפיין אותם. מחקרים דיווחו על הפחתה של 33% בהתחייבויות מוקדמות בסימולציות של ניווט רובוטי , עיקרון שניתן להשליך גם על תחום המסחר.
  + **קבלת החלטות אדפטיבית:** השימוש במשוב בייסיאני (באופן מובלע או מפורש) ובעדכון דינמי של משרעות הפעולה על סמך נתונים קונטקסטואליים (כגון מדדי תנודתיות) משפר את יכולת ההסתגלות של המערכת לטווח ארוך ולתנאי שוק משתנים.
  + **ביצועים משופרים:** במודלים של אופטימיזציית תיקים פיננסיים שבוצעו על נתוני מדד S&P 500 (לתקופה 2010-2020), QIDDM הציג ביצועים עדיפים באופן משמעותי על פני שיטות בסיס. הוא השיג תשואה שנתית של 14.3%, יחס שארפ של 1.8, ו-Maximum Drawdown של 12% בלבד. תוצאות אלו היו טובות יותר מאלו של אלגוריתם Monte Carlo Tree Search (MCTS) (תשואה 9.7%, שארפ 1.2, MDD 21%) ומאסטרטגיה חמדנית (Greedy strategy) (תשואה 6.5%, שארפ 0.7, MDD 34%).
* **חולשות:**
  + **מדרגיות (Scalability) למרחבי מצב גדולים:** ככל שמרחב המצבים (כלומר, מספר הפעולות האפשריות או המורכבות של ייצוג המצב) גדל, חישוב ועדכון המשרעות והאנטרופיה עלולים להפוך לתובעניים חישובית. הדבר דורש פיתוח של קירובים יעילים מבחינת חישוב אנטרופיה או טכניקות להפחתת מימדים.
  + **הגדרת גורמי קריסה חזקים:** קביעת סף האנטרופיה האופטימלי (\theta) והגדרת הפונקציה המדויקת (f) לעדכון המשרעות על סמך נתונים קונטקסטואליים הן משימות מאתגרות הדורשות כיול קפדני ועלולות להיות תלויות-תחום.
  + **תקפות המטאפורה הקוונטית:** חשוב להדגיש כי למרות ההשראה הקוונטית והשימוש בטרמינולוגיה קוונטית (כמו "משרעת" ו"קריסה"), QIDDM הוא אלגוריתם קלאסי הפועל על חומרה קלאסית. הוא אינו דורש מחשב קוונטי ונמנע ממגבלות החומרה והמורכבות הכרוכות במחשוב קוונטי אמיתי.

### 6.5 תובנות ביצועים ממחקר וניהול סיכונים מובנה

מנגנון "ההתחייבות המאוחרת" של QIDDM, המבוסס על "קריסת" סופרפוזיציה של פעולות פוטנציאליות רק כאשר אי-הוודאות (המיוצגת על ידי האנטרופיה) יורדת מספיק, מהווה למעשה אסטרטגיית ניהול סיכונים מובנית. המערכת נמנעת מפעולה כאשר השוק נתפס כ"לא ברור" או "רועש" (מצב המאופיין באנטרופיה גבוהה, שבו ההסתברויות לפעולות השונות דומות יחסית). לעומת זאת, היא מעדיפה לפעול כאשר יש "קונצנזוס" גבוה יותר בנתונים (מצב המאופיין באנטרופיה נמוכה, שבו פעולה אחת הופכת לסבירה הרבה יותר מהאחרות). זהו יתרון משמעותי בשווקים פיננסיים, הידועים בתקופות של אי-ודאות קיצונית ותנודתיות גבוהה.

QIDDM מטמיע עיקרון של "סבלנות אסטרטגית". הוא אינו "נאלץ" לפעול בכל צעד זמן או בכל פעם שמתקבל אות חדש. במקום זאת, הוא ממתין באופן אקטיבי ל"רגע הנכון" – הרגע שבו המידע המצטבר מצדיק התחייבות לפעולה ספציפית. גישה זו יכולה להיות יעילה במיוחד במניעת "מסחר יתר" (over-trading), תופעה שבה סוחרים מבצעים פעולות רבות מדי בתגובה ל"רעש" שוקי חסר משמעות, ובכך צוברים עלויות עסקה ומגדילים את הסיכון. כמו כן, היא יכולה לסייע במניעת כניסה לפוזיציות בתזמון גרוע עקב תגובתיות יתר לאירועים חולפים. האתגר המרכזי, כאמור, הוא לקבוע את סף האנטרופיה (\theta) בצורה אופטימלית, כך שהמערכת לא תהיה שמרנית מדי (ובכך תחמיץ הזדמנויות מסחר טובות) או אגרסיבית מדי (ותפעל מוקדם מדי על בסיס מידע לא מספק).

הטבלה הבאה מסכמת את מדדי הביצועים של QIDDM בסימולציות מסחר פיננסי, כפי שדווחו במחקר :

| מדד ביצועים | QIDDM | MCTS | אסטרטגיה חמדנית | קטע מחקר |
| --- | --- | --- | --- | --- |
| תשואה שנתית (%) | 14.3 | 9.7 | 6.5 |  |
| יחס שארפ | 1.8 | 1.2 | 0.7 |  |
| Max Drawdown (%) | 12 | 21 | 34 |  |

טבלה זו מדגישה את השיפורים הכמותיים הספציפיים ש-QIDDM מציע בהשוואה לפרדיגמות החלטה אחרות בהקשר הפיננסי שבו נבדק, ומבהירה את הצעת הערך הפוטנציאלית שלו.

## 7. ניתוח השוואתי וסינתזה

לאחר סקירת ארבע הגישות המרכזיות לשילוב סופרפוזיציה בהשראת קוונטים במודלי מסחר אלגוריתמי, ניתן לערוך ניתוח השוואתי ולזהות קווי דמיון, הבדלים, ופוטנציאל לסינרגיה.

### 7.1 השוואה צולבת של ארבע הגישות

| מאפיין | ביצוע מדיניות באנסמבל | בחירת פעולה הסתברותית | מעקב אחר השערות מצב מרובות (MHST) | התחייבות מאוחרת דמוית QIDDM |
| --- | --- | --- | --- | --- |
| **מנגנון הדמיית סופרפוזיציה** | סופרפוזיציה של מדיניויות/לוגיקות מגוונות. "קריסה" בתהליך איחוד ההחלטות. | סופרפוזיציה של פעולות אפשריות בהתפלגות. "קריסה" בדגימה מההתפלגות. | סופרפוזיציה של השערות מצב שוק. "קריסה" בהתכנסות לאמונה חזקה או גיזום השערות. | סופרפוזיציה של פעולות עם משרעות דינמיות. "קריסה" מופעלת על ידי סף אנטרופיה. |
| **דרישות חישוביות** | גבוהות, במיוחד עם סוכנים רבים/מורכבים. | תלוי במורכבות הרשת; יכול להיות בינוני. | פוטנציאל גבוה מאוד עם מספר השערות גדול. | עלול להיות גבוה עם מרחבי מצב גדולים. |
| **חוסן (Robustness)** | חוזק מרכזי; מפחית שונות ורגישות. | יכול להיות חזק אם מכויל היטב ומותאם לאי-ודאות. | פוטנציאל לחוסן גבוה בתנאי שוק משתנים/עמומים. | מתוכנן לחוסן בסביבות תנודתיות על ידי הימנעות מהתחייבויות מוקדמות. |
| **הסתגלות (Adaptability)** | יכול להסתגל עם גיוון מספק או מנגנון בחירה/שקלול דינמי. | מסתגל על ידי עדכון התפלגויות המדיניות. | הסתגלות גבוהה למשטרי שוק משתנים היא יתרון מפתח. | מסתגל על ידי עדכון משרעות הפעולה בהתבסס על נתונים קונטקסטואליים. |
| **מורכבות יישום** | יכול להיות מורכב לתכנון וכוונון (גיוון, איחוד, שקלול). | יישום יחסית פשוט עם רשתות נוירונים סטנדרטיות. | מורכב מאוד ליישום יעיל (ניהול השערות, גיזום, אסוציאציית נתונים). | דורש פיתוח מנגנון האנטרופיה והתאמת המשרעות. |
| **דגשים מדווחים על ביצועים** | יחס שארפ משופר, Max Drawdown מופחת. יציבות תשואות משופרת. | מהווה בסיס למדיניויות RL רבות. שיפור בהתמודדות עם אי-ודאות. | שיפור ביצועים בשילוב HMM עם RL לזיהוי משטרים. | תשואה שנתית 14.3%, יחס שארפ 1.8, Max Drawdown 12% (בסימולציית S&P 500). |

### 7.2 זיהוי סינרגיות וחפיפות בין גישות

ארבע הגישות אינן בהכרח בלעדיות זו לזו, וניתן לזהות פוטנציאל לסינרגיה ביניהן:

* **אנסמבל ובחירת פעולה הסתברותית:** ניתן לשלב גישות אלו בקלות. כל סוכן בודד באנסמבל יכול להפיק התפלגות הסתברויות על פני הפעולות האפשריות, ומנגנון איחוד ההחלטות של האנסמבל יכול לפעול על התפלגויות אלו (למשל, על ידי חישוב ממוצע משוקלל של ההתפלגויות או בחירת הפעולה בעלת ההסתברות הממוצעת הגבוהה ביותר).
* **MHST ו-QIDDM:** גישת MHST, ובפרט שימוש ב-HMM לזיהוי משטרי שוק, יכולה לספק מידע קונטקסטואלי רב ערך עבור מודל QIDDM. זיהוי משטר השוק הנוכחי (למשל, "תנודתיות גבוהה") יכול לשמש כקלט C לעדכון המשרעות \alpha\_i של הפעולות הפוטנציאליות ב-QIDDM, ובכך להתאים את סף ההתחייבות של QIDDM למצב השוק.
* **אנסמבל ו-QIDDM:** ניתן לבנות אנסמבל של סוכנים, כאשר כל סוכן בודד באנסמבל משתמש במנגנון פנימי דמוי QIDDM לקבלת החלטות. גישה כזו יכולה לשלב את יתרונות החוסן של אנסמבלים עם יתרונות ההתחייבות המאוחרת של QIDDM.

### 7.3 תפקיד עקרונות RL בייסיאני בניהול אי-ודאות

למידת חיזוק בייסיאנית (BRL) מספקת מסגרת מתמטית טבעית לכימות וניהול של אי-ודאות, שהיא אתגר מרכזי בכל ארבע הגישות הנדונות. הרעיון המרכזי ב-BRL הוא שמירה על התפלגויות הסתברות (אמונות) על פני פרמטרים לא ידועים של המודל (למשל, דינמיקת המעברים של השוק, פונקציית התגמול), על פני פונקציות ערך, או על פני מדיניויות. שמירה על התפלגויות אלו היא אנלוגית רעיונית להחזקת "סופרפוזיציה" של מודלים או אסטרטגיות אפשריות. ככל שמתקבל מידע חדש מהסביבה, התפלגויות אלו מתעדכנות באמצעות חוק בייס.

BRL יכול לשפר את תהליך החקירה (exploration) על ידי הכוונת הסוכן לאזורים במרחב המצב או הפעולה שבהם אי-הוודאות היא הגבוהה ביותר. הוא מאפשר שילוב של ידע מוקדם (prior knowledge) בצורה עקרונית, ויכול לסייע בהערכת סיכונים בצורה מתוחכמת יותר. תכונות אלו יכולות להועיל לכל אחת מהגישות שנסקרו:

* בגישת **האנסמבל**, ניתן לראות כל חבר אנסמבל כדגימה מהתפלגות אחורית (posterior) של מדיניויות אפשריות.
* ב**בחירת פעולה הסתברותית**, ההסתברויות עצמן יכולות להיות תוצר של מודל בייסיאני.
* ב-**MHST**, כל השערת מצב שוק יכולה להיות משויכת להתפלגות אמונה בייסיאנית, המתעדכנת עם נתונים חדשים.
* ב-**QIDDM**, עדכון המשרעות \alpha\_i וההחלטה על "קריסה" יכולים להיות מונחים על ידי עקרונות בייסיאניים של עדכון אמונות והפחתת אי-ודאות.

### 7.4 מבט מאחד: גיוון קוגניטיבי והפחתת שבירות

כל ארבע הגישות, למרות המנגנונים השונים שלהן, חולקות מטרה משותפת וחשובה: להפוך את תהליך קבלת ההחלטות במסחר ל*פחות שביר* (less brittle) ויותר *אדפטיבי* לנוכח אי-הוודאות והדינמיות המובנות של השווקים הפיננסיים. הן משיגות זאת על ידי דחיית ההתחייבות ל"אמת" יחידה (מדיניות אחת, פעולה אחת, מצב שוק אחד) מוקדם מדי, ובמקום זאת, הן חוקרות, שומרות ומעריכות מרחב של אפשרויות במקביל.

השווקים הפיננסיים הם מערכות מורכבות, רועשות ולעיתים קרובות לא-נייחות. מודלים המתחייבים במהירות רבה מדי לאסטרטגיה או לפעולה ספציפית עלולים להיות פגיעים מאוד לשינויים פתאומיים, למידע חלקי או מטעה, או להנחות יסוד שגויות לגבי מצב השוק. "ההשראה הקוונטית" לסופרפוזיציה, בהקשר זה, אינה נוגעת רק לחישוב מקביל פוטנציאלי, אלא ליתרון קונספטואלי עמוק יותר: הטמעה של *גיוון קוגניטיבי* מובנה במודל. על ידי החזקת מספר "מחשבות" (בין אם אלו מדיניויות שונות באנסמבל, פעולות אפשריות בהתפלגות, השערות מצב מרובות, או משרעות פעולה דינמיות) בו-זמנית, המודל יכול להגיב בצורה חלקה ומדודה יותר למידע חדש. הוא יכול להימנע מ"ראיית מנהרה" שבה הוא מתמקד רק באפשרות אחת, ולהיות עמיד יותר בפני טעויות קטסטרופליות הנובעות מהסתמכות על הנחה שגויה יחידה לגבי מצב העולם. זהו יתרון קונספטואלי חשוב, שעשוי להיות משמעותי אף יותר מהשיפורים הכמותיים הספציפיים בביצועים שמודלים אלו עשויים להציע.

## 8. קביעת הגישה המתאימה ביותר למסחר אלגוריתמי

### 8.1 קביעת קריטריונים ל"התאמה"

בחירת הגישה המתאימה ביותר לשילוב סופרפוזיציה בהשראת קוונטים במודל מסחר אלגוריתמי תלויה במספר קריטריונים מרכזיים, המשקפים את הדרישות המעשיות של סביבת המסחר:

1. **ביצועים מותאמי סיכון:** זהו אולי הקריטריון החשוב ביותר. הוא כולל מדדים כגון יחס שארפ גבוה (המצביע על תשואה עודפת ליחידת סיכון), Maximum Drawdown (משיכה מקסימלית) נמוך (המעיד על הגנה טובה יותר מפני הפסדים גדולים), ותשואה עקבית לאורך זמן.
2. **חוסן (Robustness):** היכולת של המודל להמשיך ולספק ביצועים טובים גם תחת תנאי שוק משתנים, תנודתיות גבוהה, או נוכחות של "רעש" בנתונים. חוסן כולל גם הפחתת רגישות לשינויים קטנים בהיפר-פרמטרים של המודל או בנתוני הקלט.
3. **הסתגלות (Adaptability):** היכולת של המודל ללמוד ולהתאים את האסטרטגיה שלו באופן דינמי בתגובה למידע חדש, לשינויים בדינמיקת השוק, או למעברים בין משטרי שוק שונים.
4. **ישימות חישובית (Computational Feasibility):** האיזון בין מורכבות המודל והפוטנציאל שלו לביצועים גבוהים, לבין המשאבים החישוביים (זמן עיבוד, כוח מחשוב, זיכרון) הנדרשים לאימונו ולהפעלתו בזמן אמת. אלגוריתמים בהשראת קוונטים, מעצם הגדרתם, מיועדים לפעול על חומרה קלאסית, אך עדיין יכולות להיות להם דרישות חישוביות משמעותיות.
5. **בשלות וזמינות של עדויות אמפיריות:** מידת הבשלות של הגישה, כפי שבאה לידי ביטוי בקיום מחקרים המדגימים יישום מוצלח שלה בהקשר של מסחר פיננסי, והזמינות של קוד פתוח או מסגרות עבודה התומכות ביישום.

### 8.2 טיעון מבוסס ראיות עבור הגישה/ות המבטיחה/ות ביותר

בהתבסס על חומר המחקר שנסקר ועל הקריטריונים שהוגדרו לעיל, ניתן להעריך את התאמתה של כל אחת מארבע הגישות:

* **שיטות אנסמבל:** גישה זו מציגה את העדויות האמפיריות החזקות והעקביות ביותר לשיפור ביצועים מותאמי סיכון וחוסן במספר רב של מחקרים. המנגנונים הבסיסיים של אנסמבלים (כגון אימון סוכנים מגוונים ואיחוד החלטותיהם) מובנים יחסית וניתנים ליישום באמצעות כלי למידת חיזוק קיימים. האתגרים העיקריים הם התקורה החישובית הכרוכה באימון מספר סוכנים ומורכבות הכוונון של האנסמבל כולו, אך היתרונות המדווחים במונחי יציבות וביצועים עשויים לעלות על חסרונות אלו עבור יישומים רבים.
* **התחייבות מאוחרת דמוית QIDDM:** גישה זו מציגה תוצאות מרשימות מאוד בסימולציות פיננסיות ספציפיות שבהן נבדקה. המנגנון האינטואיטיבי של התחייבות מאוחרת, המבוסס על סף אנטרופיה, נראה מתאים במיוחד להתמודדות עם התנודתיות הגבוהה המאפיינת שווקים פיננסיים. QIDDM נחקר פחות בהרחבה מאשר שיטות אנסמבל באופן כללי, אך התוצאות הראשוניות מבטיחות ומצביעות על פוטנציאל גבוה.
* **בחירת פעולה הסתברותית:** גישה זו מהווה רכיב בסיסי ואינהרנטי במדיניויות למידת חיזוק רבות, במיוחד אלו המבוססות על שיטות Actor-Critic או Policy Gradient. היא ניתנת לשילוב טבעי בכל אחת מהגישות האחרות. כשלעצמה, היא מספקת גמישות מובנית ויכולת לנהל אי-ודאות, אך ייתכן שלא תספק את אותה רמת חוסן או שיפור ביצועים דרמטי כמו אנסמבלים ייעודיים או QIDDM, אלא אם כן היא משולבת עם טכניקות נוספות לשיפור החוסן או ההסתגלות.
* **מעקב אחר השערות מצב מרובות (MHST):** רעיונית, גישה זו חזקה מאוד להתמודדות עם אי-ודאות לגבי משטר השוק הנוכחי או העתידי. עם זאת, יישום מלא של אלגוריתם MHT קלאסי עלול להיות מורכב מאוד מבחינה חישובית עבור הדינמיקה המהירה והרב-מימדית של שווקים פיננסיים. שילובים עם מודלי HMM או עם עקרונות של למידה בייסיאנית נראים כדרך מעשית יותר להתקדם, אך עדיין עשויים לדרוש מומחיות משמעותית בפיתוח ובכיול, וייתכן שהם בשלים פחות ליישום מיידי בהשוואה לאנסמבלים או QIDDM.

בהתבסס על הניתוח, נראה כי **שיטות אנסמבל** הן הגישה הבשלה, המבוססת והמתועדת ביותר מבין הארבע, עם עדויות ברורות לשיפור ביצועים במסחר אלגוריתמי. **אלגוריתם QIDDM** מציג פוטנציאל גבוה כגישה משלימה או חלופית, במיוחד עבור אסטרטגיות המתמקדות בהימנעות מהתחייבויות מוקדמות בתנאי תנודתיות גבוהה, וראוי למחקר נוסף ויישומים רחבים יותר.

### 8.3 הקשרת הגישה "הטובה ביותר" למטרות מסחר ספציפיות

חשוב להדגיש כי "הגישה המתאימה ביותר" אינה בהכרח גישה אחת ויחידה המתאימה לכל התרחישים. הבחירה האופטימלית עשויה להיות תלויה במטרות המסחר הספציפיות, בסובלנות לסיכון, במאפייני הנכסים הנסחרים, ובמגבלות התפעוליות והחישוביות של המשתמש. לדוגמה:

* עבור אסטרטגיות הדורשות **חוסן מקסימלי ויציבות תשואות** לאורך זמן, תוך מזעור ההשפעה של תנודות שוק קיצוניות, **שיטות אנסמבל** עשויות להיות המועדפות בשל יכולתן המוכחת להפחית שונות ולהתגבר על חולשות של סוכנים בודדים.
* עבור אסטרטגיות המתמקדות ב**תזמון כניסה ויציאה מדויק מפוזיציות בשווקים תנודתיים מאוד**, שבהם הימנעות מפעולות פזיזות היא קריטית, **אלגוריתם QIDDM** עם מנגנון ההתחייבות המאוחרת שלו עשוי להתאים במיוחד.
* במצבים שבהם **זיהוי שינויי משטר** הוא קריטי להצלחת האסטרטגיה (למשל, מעבר בין סביבת ריבית נמוכה לגבוהה, או שינוי בסנטימנט המשקיעים), גישות המשלבות **MHST (כגון HMM-RL)** עשויות להציע יתרון, גם אם הן מורכבות יותר ליישום.

יתרה מכך, במקום לחפש "כדור כסף" יחיד, ייתכן שגישה מודולרית או היברידית, המשלבת אלמנטים מכמה מהגישות הללו, תהיה היעילה ביותר. לדוגמה, ניתן להשתמש במערכת מבוססת HMM-RL לזיהוי משטר השוק ברמה גבוהה. בהתבסס על המשטר שזוהה, ניתן להפעיל אנסמבל של מדיניויות ספציפיות לאותו משטר, כאשר כל מדיניות באנסמבל משתמשת בבחירת פעולה הסתברותית או אף במנגנון פנימי דמוי QIDDM. הגמישות להתאים את הכלים והטכניקות למשימה הספציפית היא מפתח להצלחה ביישום גישות מורכבות אלו.

## 9. המלצות אסטרטגיות ומבט לעתיד

### 9.1 המלצות ישימות ליישום הגישה/ות המתאימה/ות

בהתבסס על הניתוח לעיל, ניתן להציע את ההמלצות האסטרטגיות הבאות ליישום גישות מבוססות סופרפוזיציה בהשראת קוונטים במסחר אלגוריתמי:

* **עבור יישום שיטות אנסמבל:**
  + **התחלה הדרגתית:** מומלץ להתחיל עם מספר קטן יחסית של סוכנים מגוונים (למשל, 3-5 סוכנים). ניתן להשיג גיוון על ידי שימוש באלגוריתמי RL שונים (כגון PPO, A2C, SAC), אימון על תתי-קבוצות שונות של נתונים היסטוריים, או שינוי היפר-פרמטרים מרכזיים.
  + **מנגנוני איחוד פשוטים תחילה:** יש להתמקד תחילה במנגנוני איחוד החלטות פשוטים יחסית, כגון ממוצע משוקלל של הסתברויות פעולה המבוסס על ביצועי הוולידציה (למשל, יחס שארפ), או בחירת הסוכן בעל הביצועים הטובים ביותר. ניתן להתקדם למנגנונים מורכבים יותר (כמו הצבעת רוב מתוחכמת או מודלים מבוססי מסווגים) לפי הצורך והניסיון המצטבר.
  + **תשתית אימון וולידציה מקבילית:** יש להשקיע בבניית תשתית המאפשרת אימון וולידציה מקביליים של מספר סוכנים, כפי שהודגם במחקרים המשתמשים ב-GPUs לסימולציות מסיביות. הדבר חיוני לקיצור זמני הפיתוח והבדיקה.
* **עבור יישום QIDDM או גישות דומות של התחייבות מאוחרת:**
  + **התמקדות במרחבי מצב ופעולה מוגדרים היטב:** מומלץ להתחיל עם מרחבי מצב ופעולה פשוטים יחסית כדי להקל על הכיול והבנת התנהגות המודל.
  + **הנדסת מאפיינים קונטקסטואליים משמעותיים:** יש חשיבות רבה לזיהוי ובחירה של מאפיינים קונטקסטואליים (המשתנה C במודל) המשפיעים באופן משמעותי על החלטות המסחר ועל רמת אי-הוודאות. מדדי תנודתיות, סנטימנט שוק, או אינדיקטורים מאקרו-כלכליים יכולים לשמש כנקודת התחלה.
  + **כיול זהיר של סף האנטרופיה:** סף האנטרופיה (\theta) הוא פרמטר קריטי השולט במידת ה"סבלנות" של המודל. יש לכייל אותו בקפידה באמצעות בדיקות חוזרות (backtesting) על נתונים היסטוריים מגוונים, כדי למצוא את האיזון האופטימלי בין הימנעות מהתחייבויות מוקדמות לבין ניצול הזדמנויות.

### 9.2 טיפול באתגרים כוללים ב-RL פיננסי

יישום מוצלח של גישות אלו דורש גם התייחסות לאתגרים כלליים המאפיינים למידת חיזוק בהקשר פיננסי:

* **יכולת הסבר (Explainability):** מודלי RL מורכבים, ובפרט אלו המבוססים על רשתות נוירונים עמוקות, נתפסים לעיתים קרובות כ"קופסאות שחורות", מה שמקשה על הבנת הלוגיקה מאחורי החלטותיהם ועל בניית אמון בהם. עבור אנסמבלים, ניתן לנסות ולנתח את החלטות הסוכנים הבודדים כדי לקבל תובנות. עבור QIDDM, ניתוח השינויים במשרעות הפעולה וברמת האנטרופיה לאורך זמן יכול לספק רמזים לגבי "ההיגיון" של המודל.
* **אי-נייחות (Non-stationarity):** שווקים פיננסיים הם דינמיים ומשתנים ללא הרף, מה שהופך את ההנחה על סביבה נייחת (שעליה מתבססים אלגוריתמי RL רבים) לבעייתית. כל הגישות שנדונו מנסות להתמודד עם אי-נייחות בדרכים שונות (למשל, גיוון והתאמה באנסמבלים, עדכון משרעות דינמי ב-QIDDM, זיהוי משטרים ב-MHST). אימון מתמשך (continuous training), ולידציה תקופתית על נתונים עדכניים, ומנגנוני הסתגלות מהירים הם חיוניים.
* **מידול MDP (בחירת מצבים, פונקציות תגמול):** ההגדרה של מרחב המצבים (state space), מרחב הפעולות (action space), ופונקציית התגמול (reward function) היא קריטית להצלחת כל יישום RL. יש להתאים את עיצוב רכיבים אלו למאפיינים הספציפיים של כל אחת מהגישות המבוססות על סופרפוזיציה, ולוודא שהם משקפים נאמנה את מטרות המסחר ואת סובלנות הסיכון.

### 9.3 כיווני מחקר עתידיים פוטנציאליים

התחום של שילוב עקרונות בהשראת קוונטים במסחר אלגוריתמי מבוסס RL הוא עדיין בראשית דרכו, וקיימים מספר כיווני מחקר עתידיים מבטיחים:

* **שילוב עמוק יותר של למידה בייסיאנית:** הרחבת השימוש בעקרונות בייסיאניים בכל ארבע הגישות יכולה לשפר משמעותית את כימות אי-הוודאות, את ניהול הסיכונים, ואת יכולת ההסתגלות של המודלים.
* **מנגנוני איחוד החלטות מתוחכמים לאנסמבלים:** פיתוח מנגנוני איחוד החלטות דינמיים ואדפטיביים יותר עבור אנסמבלים, אולי בהשראת תיאוריות של קבלת החלטות קבוצתית, למידת חיזוק מרובת סוכנים (MARL) , או טכניקות למידה המבוססות על תשומת לב (attention mechanisms).
* **הרחבת QIDDM:** חקירת יישומים של QIDDM במרחבי פעולה רציפים (continuous action spaces) או מורכבים יותר, ופיתוח שיטות לכיול אוטומטי של סף האנטרופיה והמשקלים הקונטקסטואליים.
* **שיטות MHST יעילות חישובית:** פיתוח שיטות MHST יעילות יותר מבחינה חישובית עבור שווקים פיננסיים, אולי תוך שימוש בטכניקות קירוב, דגימה חכמה, או שימוש ברשתות נוירונים עמוקות להטמעת השערות (embedding of hypotheses) במרחב לטנטי נמוך-מימד, כפי שהוצע ב-.
* **חקירת "השראה קוונטית אמיתית" מעבר לאנלוגיות:** בחינה של שימוש באלמנטים מתקדמים יותר מאלגוריתמים בהשראת קוונטים, כגון שימוש ברשתות טנזוריות (Tensor Networks) לייצוג יעיל של מצבים מורכבים , או שילוב של אלגוריתמי אופטימיזציה בהשראת קוונטים (כגון גרסאות קלאסיות של QAOA) בתוך מסגרות ה-RL הקיימות.

הפוטנציאל האמיתי בתחום זה עשוי להיות לא ביישום "טהור" של אחת הגישות שנסקרו, אלא ב*הכלאה חכמה* (intelligent hybridization) שלהן, הממנפת את החוזקות של כל אחת תוך מזעור חולשותיה. יתר על כן, שילוב גישות אלו עם התקדמויות אחרות בתחום הבינה המלאכותית, כגון מודלי שפה גדולים (LLMs) המסוגלים לעבד ולהבין מידע טקסטואלי (כגון חדשות פיננסיות או דוחות אנליסטים) וסנטימנט שוק , יכול לספק קלט קונטקסטואלי עשיר ומדויק יותר למודלי המסחר. לדוגמה, מערכת MHST יכולה לזהות את משטר השוק הנוכחי. בהתבסס על המשטר שזוהה, ניתן לבחור אנסמבל ספציפי של סוכני QIDDM, כאשר כל סוכן QIDDM מקבל קלט לא רק מנתוני מחירים ואינדיקטורים טכניים, אלא גם מניתוח סנטימנט בזמן אמת המתקבל ממודל LLM. ארכיטקטורה היברידית כזו עשויה להוביל לדור הבא של מודלי מסחר אלגוריתמי, המסוגלים להתמודד בצורה הוליסטית ומתוחכמת יותר עם המורכבות והדינמיות של השווקים הפיננסיים המודרניים.

#### Works cited

1. Algorithmic Trading: Definition, How It Works, Pros & Cons - Investopedia, https://www.investopedia.com/terms/a/algorithmictrading.asp 2. What is Algo Trading? Difference, Benefits, Strategies & Regulations - Groww, https://groww.in/blog/what-is-algo-trading 3. Revisiting Ensemble Methods for Stock Trading and Crypto Trading Tasks at ACM ICAIF FinRL Contests 2023/2024 - arXiv, https://arxiv.org/html/2501.10709v1 4. arxiv.org, https://arxiv.org/pdf/2501.17992 5. What are quantum-inspired algorithms, and how do they differ from true quantum algorithms? - Milvus, https://milvus.io/ai-quick-reference/what-are-quantuminspired-algorithms-and-how-do-they-differ-from-true-quantum-algorithms 6. Quantum-Inspired Algorithms - GeeksforGeeks, https://www.geeksforgeeks.org/quantum-inspired-algorithms/ 7. A Survey of Quantum Computing for Finance - arXiv, https://arxiv.org/pdf/2201.02773 8. Ensemble Models in ML: Techniques and Benefits - SoluLab, https://www.solulab.com/ensemble-modeling/ 9. openfin.engineering.columbia.edu, https://openfin.engineering.columbia.edu/sites/default/files/content/publications/ensemble.pdf 10. Multi-Agent Reinforcement Learning with Focal Diversity Optimization - arXiv, https://arxiv.org/html/2502.04492v1 11. (PDF) A Deep Ensemble Reinforcement Learning Based Approach ..., https://www.researchgate.net/publication/387921640\_A\_Deep\_Ensemble\_Reinforcement\_Learning\_Based\_Approach\_For\_Stock\_Trading 12. Ensemble Learning Techniques: Boosting, Bagging, and Stacking Explained, https://www.interviewnode.com/post/ensemble-learning-techniques-boosting-bagging-and-stacking-explained 13. Multi-Agent Reinforcement Learning: A Review of Challenges and Applications - MDPI, https://www.mdpi.com/2076-3417/11/11/4948 14. Reinforcement Learning For Trading Applications - Alpha Architect, https://alphaarchitect.com/reinforcement-learning-for-trading/ 15. Probabilistic Constrained Reinforcement Learning with Formal Interpretability - arXiv, https://arxiv.org/html/2307.07084v4 16. Probabilistic price forecasts for short-term trade optimization - Dexter ..., https://dexterenergy.ai/news/probabilistic-price-forecasts-for-short-term-trade-optimization/ 17. Game Theory and Multi-Agent Reinforcement Learning : From Nash Equilibria to Evolutionary Dynamics - arXiv, https://arxiv.org/html/2412.20523v1 18. www.cs.toronto.edu, https://www.cs.toronto.edu/~cebly/Papers/\_download\_/bayesMARL.pdf 19. General Multi Hypotheses tracking implementation example - Stone Soup's documentation!, https://stonesoup.readthedocs.io/en/v1.4/auto\_examples/dataassociation/mht\_example.html 20. Reinforcement Learning with Hidden Markov Models for Discovering Decision-Making Dynamics - arXiv, https://arxiv.org/html/2401.13929v1 21. www.cloud-conf.net, https://www.cloud-conf.net/datasec/2025/proceedings/pdfs/IDS2025-3SVVEmiJ6JbFRviTl4Otnv/966100a067/966100a067.pdf 22. citeseerx.ist.psu.edu, https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=064dec29989aafd3b95aa0eff261556ba70dd363 23. Quantum-Inspired Dynamic Decision-Making Algorithm (QIDDM): A Robust Framework for Delayed Commitment in Uncertain Environments - ResearchGate, https://www.researchgate.net/publication/390338061\_Quantum-Inspired\_Dynamic\_Decision-Making\_Algorithm\_QIDDM\_A\_Robust\_Framework\_for\_Delayed\_Commitment\_in\_Uncertain\_Environments 24. www.ijraset.com, https://www.ijraset.com/best-journal/quantum-inspired-dynamic-decision-making-algorithm-qiddm-a-robust-framework-for-delayed-commitment-in-uncertain-environments 25. The Evolution of Reinforcement Learning in Quantitative Finance: A Survey - arXiv, https://arxiv.org/html/2408.10932v3 26. Multi-agent Reinforcement Learning: A Comprehensive Survey - arXiv, https://arxiv.org/html/2312.10256v2 27. SPD: Synergy Pattern Diversifying Oriented Unsupervised Multi-agent Reinforcement Learning, https://proceedings.neurips.cc/paper\_files/paper/2022/file/825341ab91db01bf063add41ac022702-Paper-Conference.pdf 28. arXiv:2502.11433v3 [cs.AI] 19 Feb 2025, https://arxiv.org/pdf/2502.11433