الجمهورية العربية السورية
وزارة التعليم العالي والبحث العلمي
المعهد العالي لإدارة الأعمال
قسم الإدارة المالية والمصرفية

أثر الأسلوب المستخدم لبناء محفظة الأسهم في أدائها
(دراسة تطبيقية على أسهم سوقي دمشق وعمّان للأوراق المالية)

The effect of stock portfolio creation method on its performance
(An empirical study on Damascus Stock Exchange and Amman Stock Exchange)

رسالة أعدت لنيل درجة الماجستير في علوم الإدارة

اختصاص الإدارة المالية والمصرفية

إعداد الباحث

وسيم حبي

إشراف

د. راغب الغصين

العام الدراسي 2021-2022

((لا يعبر هذا العمل إلا عن وجهة نظر معده، ولا يتحمل المعهد أية مسؤولية جراء هذا العمل))

الملخص

تهدف هذه الدراسة إلى استخدام التعلم المعزّز العميق وخوارزميات التداول المبنية على كل من استراتيجيتي الزخم والمعاكسة لإدارة ثلاثة أنواع من المحافظ الاستثمارية المستمرة ومقارنة أداءها. تتكون المحافظ من أسهم منتقاة من سوق دمشق للأوراق المالية باستخدام مجموعة من المعايير. حيث تم جمع أسعار الأسهم عن الفترة من 2011/01/09 إلى 2020/07/30. كما تم إعادة الدراسة في بورصة عمّان وجمع أسعار الأسهم عن الفترة من 2011/01/09 إلى 2020/07/30.

لإجراء المقارنة تم استخدام معايير العائد السنوي، الانحراف المعياري للعوائد، قيمة المحفظة التراكمية، نسبة شارب، ونسبة الخسارة العظمى. ولاختبار المعنوية الإحصائية للفروقات بين أزواج المحافظ الثلاث تم استخدام اختبار (Opdyke, 2007) الإحصائي المخصص لنسبة شارب. حيث أظهرت النتائج تفوق محفظة نموذج التعلم المعزّز العميق على محفظتي الزخم والمعاكسة في كل من سوق دمشق وبورصة عمّان. وتم تفسيرها بقدرة نموذج التعلم المعزّز العميق على تعلم الأنماط المختلفة من البيانات التاريخية واستخدامها بشكل فعّال لتحقيق العوائد المرتفعة. كما أظهرت نتائج المقارنة بين أداء محفظة الزخم ومحفظة المعاكسة تفوق محفظة الزخم على محفظة المعاكسة وذلك في كل من السوقين، الأمر الذي يعزى إلى فترة التداول القصيرة التي تعطي الأفضلية لاستراتيجية الزخم.

توصي الدراسة باستخدام نموذج التعلم المعزّز العميق، واستخدام خوارزميات الزخم والابتعاد عن خوارزميات المعاكسة في الأجل القصير. كما توصي المستمرين غير المحبذين للمخاطر باستخدام خوارزمية الزخم ONS التي أظهرت مؤشرات خطر منخفضة.

الكلمات المفتاحية: إدارة المحفظة المستمرة، التعلم الآلي، الشبكات العصبية الاصطناعية، التعلم المعزّز، الزخم، المعاكسة، سوق دمشق للأوراق المالية، بورصة عمّان.

Abstract

This study aims to use Deep Reinforcement Learning, along with Momentum and Contrarian trading algorithms to manage three types of Online Portfolios and compare their performance. The portfolios contain stocks from Damascus Stock Exchange (DSE) which were picked using a set of metrics. Stock prices were collected from 09/01/2011 to 30/07/2020. The study was also carried out on Amman Stock Exchange (ASE), and the stock prices were collected from 09/01/2011 to 30/07/2020.

To compare the portfolios' performance, these metrics were used: annual return, standard deviation of returns, accumulated portfolio value, Sharpe ratio, and maximum drawdown ratio. To test the statistical significance of the differences between each pair of the three portfolios, (Opdyke, 2007) test was used, which is specifically designed to be used with Sharpe ratios. The results show that the DRL portfolio outperforms both the Momentum and the Contrarian portfolios in both DSE and ASE. This was explained by the DRL model's ability to learn the trading patterns from the historical data and use it effectively to achieve high returns. The comparison results also show that the Momentum portfolio outperforms the Contrarian portfolio in both markets. And it is explained by the short trading period which is more suitable for the Momentum strategy.

The study recommends using the DRL model, and also using the Momentum algorithms and avoiding the Contrarian algorithms in the short run. It also advises the risk-averse investors to use the ONS Momentum algorithm which showed minimal risk metrics.

Keywords: Online Portfolio Selection**,** Machine Learning**,** Artificial Neural Networks, Reinforcement Learning**,** Momentum, Contrarian, Damascus Stock Exchange (DSE), Amman Stock Exchange (ASE)

الشكر والتقدير

أوجه كل الشكر والتقدير لكل من ساهم في إنجاز هذا العمل

الدكتور راغب الغصين الذي تفضّل بالإشراف على هذه الرسالة

ومنحني الوقت والصبر طيلة فترة الإعداد

السادة أعضاء لجنة الحكم الموقرة، الدكتورة منال الموصلي والدكتور ياسر كفا

لكم فائق الاحترام والتقدير

فهرس المحتويات

الملخص ..... A

الفصل التمهيدي – الإطار العام للبحث 1

مقدمة 1

1-1- الدراسات السابقة 3

ما يميز هذه الدراسة 9

1-2- مشكلة البحث 9

1-3- فرضيات البحث 10

1-4- أهمية البحث 10

1-5- أهداف البحث 11

1-6- منهج البحث 11

1-7- مجتمع وعينة البحث 12

الفصل الأول – المحافظ الاستثمارية 13

مقدمة 13

2-1- المبحث الأول – أساسيات المحافظ الاستثمارية 14

2-1-1- مفهوم المحافظ الاستثمارية 14

2-1-2- أهمية المحافظ الاستثمارية 14

2-1-3- وظائف المحافظ الاستثمارية 15

2-1-4- أنواع المحافظ الاستثمارية 17

2-1-5- التنويع في المحافظ الاستثمارية 18

2-1-6- أساسيات تكوين المحافظ الاستثمارية 20

2-2- المبحث الثاني – إدارة المحفظة المستمرة Online Portfolio Selection 22

تمهيد 22

2-2-1- التحديات التي تواجه طرق الاستثمار التقليدية 23
2-2-2- التداول الخوارزمي 24
2-2-3- التمثيل الرياضي لمشكلة إدارة المحفظة المستمرة 25
2-2-4- خوارزميات إدارة المحفظة المستمرة 27
2-2-5- خوارزميات استراتيجية الزخم 29
2-2-6- خوارزميات استراتيجية المعاكسة 31
الفصل الثاني – التعلم الآلي 33
مقدمة 33
3-1- المبحث الأول – الشبكات العصبية الاصطناعية Artificial Neural Networks 34
تمهيد 34
3-1-1- تعريف الشبكات العصبية الاصطناعية 35
3-1-2- نشأة وتطور الشبكات العصبية الاصطناعية 35
3-1-3- بنية والية عمل عصبون البيرسيبترون 38
3-1-4- بنية الشبكات العصبية الاصطناعية 40
3-1-5- تدريب الشبكات العصبية الاصطناعية 41
3-2- المبحث الثاني – التعلم المعزّز Reinforcement Learning 45
تمهيد 45
3-2-1- تطبيقات التعلم المعزّز 46
3-2-2- صفات التعلم المعزّز 46
3-2-3- التطور التاريخي للتعلم المعزّز 47
3-2-4- عمليات اتخاذ القرار الماركوفية Markov Decision Processes 48
3-2-5- خوارزميات التعلم المعزّز 50
3-2-6- الية تدريب عميل التعلم المعزّز 51
3-2-7- التعلم المعزّز العميق Deep Reinforcement Learning 53

3-2-8- معالجة مشكلة إدارة المحفظة المستمرة باستخدام التعلم المعزّز العميق.....54
3-2-9- معلمات النموذج الكلي Hyperparameters.....58
الفصل الرابع – الدراسة التطبيقية.....60
مقدمة 60
4-1- المبحث الأول – لمحة عن الأسواق المالية في الدراسة 61
تمهيد 61
4-1-1- لمحة عن سوق دمشق للأوراق المالية 61
4-1-2- توزع الشركات المدرجة في دمشق للأوراق المالية 62
4-1-3- لمحة عن بورصة عمّان 62
4-1-4- أهداف بورصة عمّان 63
4-2- المبحث الثاني – مقارنة أداء الخوارزميات في كل من سوقي دمشق وعمّان 64
تمهيد 64
4-2-1- جمع البيانات واختيار أسهم الدراسة 64
4-2-2- فترات البيانات.....68
4-2-3- معايير تقييم المحافظ.....69
4-2-4- نتائج النماذج المطبقة على سوق دمشق 70
4-2-5- نتائج النماذج المطبقة على بورصة عمّان.....76
4-3- المبحث الثالث – اختبار الفرضيات والنتائج والتوصيات 82
تمهيد 82
4-3-1- اختبار الفرضيات.....82
4-3-2- النتائج 86
4-3-3- التوصيات.....87
4-3-4- افاق البحث المستقبلية 88

قائمة الاشكال

الشكل 1 – عصبون البيرسيبترون 38
الشكل 2 – بنية الشبكة العصبية 40
الشكل 3 – التفاعل بين العميل والبيئة 51
الشكل 4 – قيمة المحافظ التراكمية للخوارزميات المطبقة في سوق دمشق على بيانات الاختبار 72
الشكل 5 – متوسط أسعار الأسهم التي تم اختيارها من سوق دمشق.....73
الشكل 6 – تركيبة محفظة نموذج DRL without TC في سوق دمشق على بيانات الاختبار 74
الشكل 7 – تركيبة محفظة نموذج DRL–TC في سوق دمشق على بيانات الاختبار.....75
الشكل 8 – قيمة المحافظ التراكمية للخوارزميات المطبقة في بورصة عمّان على بيانات الاختبار 78
الشكل 9 – متوسط أسعار الأسهم التي تم اختيارها من بورصة عمّان 78
الشكل 10 – تركيبة محفظة نموذج DRL without TC في بورصة عمّان على بيانات الاختبار 80
الشكل 11 – تركيبة محفظة نموذج DRL–TC في بورصة عمّان على بيانات الاختبار 80

قائمة الجداول

الجدول 1 – اختيار اسهم سوق دمشق للأوراق المالية 66
الجدول 2 – اختيار اسهم بورصة عمّان للأوراق المالية.....67
الجدول 3 – مجموعات بيانات الدراسة لكل سوق مالي 68
الجدول 4 – نتائج خوارزميات التداول في سوق دمشق للأوراق المالية على بيانات الاختبار 70
الجدول 5 – نتائج خوارزميات التداول في بورصة عمّان على بيانات الاختبار 76
الجدول 6 – نسبة شارب لكل نوع من المحافظ الثلاث في كل من سوق دمشق وبورصة عمّان 83
الجدول 7 – نتائج اختبار الفروقات الاحصائي لنسب شارب في كل من سوق دمشق وسوق عمّان 84

الفصل التمهيدي – الإطار العام للبحث

مقدمة

تعد إدارة المحافظ الاستثمارية أحد فروع الإدارة المالية. وتهتم بتوزيع الأموال المتاحة للاستثمار على عدد من الأصول بشكل يعظم قيمة المحفظة في أجل لاحق. بداية الأبحاث العلمية في هذا المجال تعود إلى نظرية المحفظة الحديثة التي تعظم قيمة محفظة من خلال المقايضة بين العائد المتوقع والخطر.

تركز نظرية نمو رأس المال التي جاءت فيما بعد على تعظيم معدل نمو عائد محفظة أوراق مالية أو لوغاريتم العائد المتوقع. كلا النظريتين السابقتين تصلحا لحل مشكلة إدارة المحفظة، إلا أن النظرية اللاحقة تعد أكثر ملائمة لمعالجة مشكلة إدارة المحفظة المستمرة Online Portfolio Selection المعنية بإعادة هيكلة المحفظة على عدة فترات خلال عمرها. فهي تركّز على فكرة التسلسلية أي ان عملية إعادة الهيكلة يمكن أن تحصل بشكل يومي، أسبوعي، شهري أو سنوي.

خلال العقود الأخيرة ظهرت العديد من الخوارزميات التي تقوم بمعالجة إدارة المحفظة المستمرة. بعض هذه الخوارزميات مبنية على إستراتيجيات تستغل ظواهر معروفة في الأسواق المالية، كظاهرة الزخم وظاهرة المعاكسة. وبعضها الآخر يقوم بتسخير آخر التطورات في مجال التعلم الآلي.

يعد التعلم الآلي أحد فروع الذكاء الاصطناعي ويهتم بحل مشاكل كالتصنيف والتنبؤ من خلال تدريب نموذج على عينة من البيانات دون برمجة حل المشكلة بشكل مباشر. أحد انواع التعلم الآلي هو التعلم الموجّه والذي يتضمن عدة خوارزميات أهمها الشبكات العصبية الاصطناعية Artificial Neural Networks، وهي مستوحاة من الخلايا العصبية البيولوجية، حيث تتألف من اجزاء متصلة تدعى عصبونات. تأتي قوة الشبكات العصبية من قدرتها على حساب أي دالة بشكل تقريبي مهما كانت درجة تعقيدها.

التعلم المعزّز Reinforcement Learning هو نوع آخر من أنواع خوارزميات التعلم الآلي، حيث يقوم على تدريب عميل على تعظيم قيمة مكافأة في بيئة تعكس أوضـــاع المشـــكلة التي تحاول الخوارزمية معالجتها، حيث يتم نمذجة الحالات المختلفة التي يمكن أن تواجه العميل، وتتم عملية التدريب من خلال اتخاذ قرارات بشـــكل متتابع بحيث يصـــل العميل في نهاية عملية التعلم إلى أفضـــل القرارات التي يمكن اتخاذها في أي وضع من أوضاع البيئة.

يمكن دمج قدرة الشـــبكات العصـــبية العميقة على تمثيل الملامح وفعالية التعلم المعزّز في البحث عن اســـتراتيجية مثلى لتشـــكيل ما يعرف بخوارزميات التعلم المعزّز العميق Deep Reinforcement Learning (او DRL اختصارا) والتي أظهرت نتائج مذهلة في مجال إدارة المحافظ الاستثمارية المستمرة.

نظراً للحداثة النســـبية لهذه الخوارزمية وأهميتها العملية، وبهدف معرفة مكانها بين خوارزميات التداول التقليدية المبنية على اسـتراتيجيات الزخم والمناقضـة، يأتي دور البحث في دراسـة أثر اسـتخدامها في أداء المحافظ المالية، والذي سيتم تقييمه من خلال عدد من المؤشرات المالية.

1-1- الدراسات السابقة

استخدمت الدراسات السابقة شكل من اشكال خوارزمية DRL لإدارة محفظة مكونة من أسهم أو عملات الكترونية، كما تضمن أغلبها استخدام واحدة أو أكثر من خوارزميات استراتيجية الزخم أو استراتيجية المعاكسة أو الاثنين معاً، حيث تم اجراء مقارنة في جميع الدراسات لمعرفة اداء المحافظ المشكلة حسب كل نوع من الخوارزميات. من المعايير المستخدمة في المقارنة: قيمة المحفظة التراكمية، نسبة شارب ونسبة الخسارة العظمى. فيما يلي الدراسات مرتبة حسب التسلسل الزمني:

1) دراسة (Jiang et al., 2017) بعنوان:

A Deep Reinforcement Learning Framework for the Financial Portfolio Management Problem

طبق البحث خوارزمية التعلم المعزّز العميق بهدف بناء نموذج لإدارة محفظة مالية في سوق العملات الالكترونية. يُمثَّل الوضع state في بيئة العميل في كل فترة من خلال أسعار العملات، بالإضافة إلى أوزان المحفظة في الفترة السابقة وذلك لدفع العميل إلى أخذها بعين الاعتبار من أجل إبقاء مصاريف التداول بحدود منخفضة. وتتكون مكافأة العميل في كل فترة من عائد المحفظة المحقق في نهاية الفترة والذي يتم تخفيضه بمصاريف التداول بشكل يتناسب مع التغير في أوزان المحفظة بين الفترة الحالية والسابقة. وتتألف مجموعة القرارات التي يمكن للعميل اتخاذها من أوزان العملات المختلفة ضمن المحفظة، يتم اختيار هذه الأوزان من خلال شبكة عصبية. حيث تمت تجربة ثلاثة انواع من الشبكات وهي RNN، LSTM، CNN.

استخدم الباحث بيانات أسعار نصف ساعية بين عامي 2014 و2017، وقام بتقسيمها إلى فترات تدريب، فترات اختبار، بالإضافة إلى فترة تحقق لضبط معالم النماذج. واعتمد على معايير قيمة المحفظة التراكمية، نسبة شارب ونسبة الخسارة العظمى MDD، وذلك لمقارنة نماذج البحث الثلاث مع عدد من الخوارزميات الأخرى.

تتضمن الخوارزميات التي استخدمت للمقارنة ONS، EG وUP التابعة لاستراتيجية الزخم، وخوارزميات Anticore، OLMAR، PAMR وغيرها التابعة لاستراتيجية المعاكسة، بالإضافة إلى بعض خوارزميات خط الأساس وخوارزميات مطابقة النمط.

أعطت نماذج DRL نتائج إيجابية عالية، فقد تفوقت على جميع الخوارزميات الأخرى وحققت نمو بقيمة المحفظة بمقدار 4 أضعاف المبلغ المستثمر وخصوصاً النموذج الذي استخدم الشبكة العصبية من نوع CNN، وكان متوسط قيم محافظ استراتيجية المعاكسة في نهاية فترة الاختبار مشابها لمتوسط قيم محافظ الزخم إلا أن نسبة الخسارة العظمى لمحافظ المعاكسة كانت أعلى ب 3.5 مرة من نسبة الخسارة العظمى لمحافظ الزخم. أحد الأسباب التي أدت إلى هذا الفرق في الأداء بين نماذج DRL والنماذج الأخرى هو معدل إعادة هيكلة المحفظة المرتفع (بشكل نصف ساعي)، الذي نتج عنه مصاريف تداول مرتفعة ولم تكن النماذج الأخرى قادرة على التكيف مع هذا الأمر على عكس نماذج DRL.

2) دراسة (Guo et al., 2018) بعنوان:

Robust Log-Optimal Strategy with Reinforcement Learning

تهدف هذه الدراسة إلى تطوير خوارزميتين لإدارة المحافظ الاستثمارية، الأولى هي من نوع مطابقة النمط والثانية من نوع Meta Learning، ومن ثم تقوم بمقارنتهما مع استراتيجيات الزخم والمعاكسة.

البيانات المستخدمة هي أسعار 100 سهم من الأسهم المكونة لمؤشر CSI300 الصيني. اعتمدت استراتيجية الزخم على اختيار أفضل سهم في اليوم السابق (والمعاكسة اسوأ سهم) واستثمار كامل قيمة المحفظة فيه، مما يعني غياب التنويع في كلا الاستراتيجيتين.

تمت المقارنة باستخدام معيار القيمة التراكمية للمحفظة وكانت النتائج إيجابية لصالح الخوارزمية الثانية التي استخدمت DRL وجاءت بعدها الخوارزمية الأولى لوحدها، ومن ثم استراتيجية الزخم واخيراً استراتيجية المعاكسة التي حققت نتائج سلبية في بعض الاختبارات.

3) دراسة (Weng et al., 2020) بعنوان:

Portfolio Trading System of Digital Currencies: A Deep Reinforcement Learning with Multidimensional Attention Gating Mechanism.

في هذه الدراســة يقوم النموذج المســتخدم على تدريب عميل التعلم المعزّز على اتخاذ قرار إعادة الهيكلة مباشرة من خلال شبكة عصبية تعطي كمخرجات لها أوزان المحفظة مباشرة دون التنبؤ بأسعار الأصول. تم اســتخدام عدد كبير من الأصــول في المحفظة حيث تضــمنت 20 عملة الكترونية لتدريب واختبار النموذج.

أجريت عدة اختبارات على فترات مختلفة بين 2016/06/04 و 2018/02/01 كان فيها سوق العملات الالكترونية صــاعداً، مســتقراً وهابطاً، حيث اســتخدمت المعايير التقليدية لتقييم النموذج ومقارنته مع خوارزميات أخرى منها خوارزميتي PAMR وCWMR المبنيتين على إستراتيجية المعاكسة.

كانت عوائد خوارزميتي المعاكســة ايجابية في الفترة الصــاعدة، لكنها حققت خســائر في الفترات الأخرى. بالمقارنة، حقق نموذج البحث الذي وظف التعلم المعزّز العميق عوائد ايجابية في جميع الفترات، كما كانت نسبة خسارته العظمى اقل من نسبة الخسارة العظمى لخوارزميتي المعاكسة بشكل كبير.

4) دراسة (Gao et al., 2020) بعنوان:

Application of Deep Q-Network in Portfolio Management

تســتخدم هذه الدراســة Deep Q-Network وهي إحدى خوارزميات التعلم المعزّز العميق التي توظف Q-learning كخوارزمية التعلم المعزّز بالإضــافة إلى شــبكة عصــبية، حيث تتميز عن اســتخدام Q-learning بمفردها بأنه لا يوجد شــرط بأن تكون الأوضــاع محددة أو ذات نهاية، لكنها تتطلب أن تكون قرارات العميل محددة، لذلك قام الباحث بتقسيم المحفظة إلى N جزء ومن ثم عرف مجموعة القرارات على أنها التركيبات الممكنة من أوزان M سهم حيث أن كل سهم يمكن أن يأخذ جزء واحد أو أكثر ضمن

المحفظـة. يوظف البـاحـث طريقـة Prioritized Experience Replay التي ترفع قـدرة العميـل على التدريب من خلال تركيز التعلم على العينات التي تعكس الأحداث المهمة في البيانات.

تم اختبار هذه الخوارزمية بالإضـافة إلى خوارزميات اسـتراتيجيتي الزخم والمعاكسـة على بيانات اسـعار 5 أسـهم شـركات اميركية واسـتخدمت معايير العائد التراكمي، نسـبة شـارب ونسـبة الخسـارة العظمى للمقارنة. حيث تم تدريب واختبار النموذج في الفترة بين 2015/01/02 و 2017/11/17.

تفوقت خوارزمية البحث على الخوارزميات الاخرى في فترة الاختبار، إلا أن أدائها كان معادلاً لأداء باقي الخوارزميات في البداية، وكانت خوارزميات المعاكسة أفضل من خوارزميات الزخم وسطياً.

5) دراسة (Zhang et al., 2020) بعنوان:

Cost-Sensitive Portfolio Selection via Deep Reinforcement Learning

تســعى هذه الدراســة إلى تطوير خوارزمية باســتخدام التعلم المعزّز العميق لحل مشــكلة إدارة المحافظ الاســتثمارية، ومن ثم مقارنة فعاليتها مع خوارزميات اخرى تابعة لعدد من الاســتراتيجيات من ضــمنها استراتيجيتي الزخم والمعاكسـة. حيث استخدمت الارتباط بين الأصـول كعامل لإدارة المحفظة وأخذت بعين الاعتبار مخاطر التداول ومصاريف التداول معاً.

معايير التقييم المستخدمة هي قيمة المحفظة التراكمية، نسبة شارب، نسبة كالمار Calmar Ratio (التي تحسب من خلال قسمة نسبة شارب على نسبة الخسارة العظمى) ومعدل دوران المحفظة. تم اجراء المقارنة على 4 مجموعات من بيانات العملات الإلكترونية تتكون كل مجموعة من عدد عملات (تتراوح بين 12 و44 عملة) وفترات تدريب واختبار تتراوح بين 2016/01 و 2018/08. حققت خوارزمية التعلم المعزّز العميق أفضـل النتائج من حيث جميع المعايير وكانت خوارزميات المعاكسـة أفضـل من خوارزميات الزخم وسـطيا. أعاد الباحث المقارنة على بيانات أسـهم مؤشـر SP500 فأعطت نتائج مشـابهة لبيانات العملات الإلكترونية.

6) دراسة (Luo et al., 2020) بعنوان:

A Framework of Deep Reinforcement Learning for Stock Evaluation Functions

طبق الباحث نموذج DRL على عدد من الأسـهم التايوانية، تتكون المحفظة في هذه الدراسـة من سـهم واحد، من المفترض ان عميل التعلم المعزّز بعد تدريبه سـيكون قادرا على تمييز الاوقات الملائمة للاستثمار في هذا السهم والاوقات التي يكون فيها الخطر مرتفعاً التي يجب أن يتخلى فيها عن السهم من خلال إبقاء قيمته كأموال مجمدة، وذلك من خلال تدريبه على البيانات التاريخية التي تتضـمن أسـعار وحجوم تداول السهم.

تم اختيار عدد من الأسـهم التي تعد الأكثر تمثيلاً لعدد من القطاعات لاسـتخدامها في النموذج (بشـكل فردي). كانت مدة فترة التدريب تسـع سـنوات من 2009/7/31 وحتى 2018/7/31، أما فترة الاختبار فكانت سـنة واحدة من 2018/8/1 وحتى 2019/8/1. تم إجراء مقارنة مع خوارزمية OLMAR التابعة لاستراتيجية المعاكسة باستخدام المعايير التقليدية بالإضافة إلى نسبة مجموع الأرباح إلى مجموع الخسائر التي تعبر عما تجنيه الخوارزمية لكل وحدة نقدية من الخسارة.

كانت النتائج جيدة ولصـالح خوارزمية DRL لكن لم يكن هنالك فرق جوهري بينها وبين خوارزمية OLMAR من حيث العوائد، بالإضافة إلى ذلك، لم يكن هنالك فرق معنوي إحصائياً بين نسبة الخسارة العظمى لكلا الخوارزميتين. تشـير هذه النتائج إلى عدم فعالية المحفظة المكونة من سـهم واحد بسـبب ابتعادها عن مبدأ التنويع.

7) دراسة (Huang and Tanaka, 2022) بعنوان:

MSPM: A modularized and scalable multiagent reinforcement learning-based system for financial portfolio management

تهدف هذه الدراسة إلى معالجة كل من قابلية التوسع من جهة وإمكانية إعادة الاستخدام من جهة أخرى لنماذج التعلم المعزّز في إدارة المحافظ المالية.

يستخدم نموذج البحث نوعين من الوحدات النمطية Modules القابلة للاستخدام في حالات مختلفة. يتكون النوع الأول من شبكة عصبية عميقة تقوم بالتدريب على بيانات الأسعار التاريخية بالإضافة إلى البيانات السلوكية لنوع واحد من الأصول بهدف توليد إشارات الشراء، البيع او التخطي في كل فترة زمنية.

أما النوع الثاني فهو مسؤول عن إدارة المحفظة ككل من خلال استخدام مخرجات الوحدات النمطية من النوع الأول المتصلة معه والتي تمثل مجتمعة بالإضافة لبيانات الأسعار التاريخية اوضاع البيئة المختلفة التي سيقوم العميل باستكشافها وتعلم السياسة المثلى منها التي ستتيح الوصول إلى الأوزان التي تعظم عوائد المحفظة. إن تنظيم هذه المهام على شكل وحدات نمطية منفصلة يتيح قابلية تدريب النموذج على عدد من الأصول وإضافة أصول أخرى في وقت لاحق دون الحاجة إلى إعادة تدريب النموذج كاملاً.

تم اختبار النموذج على محفظتين مكونتين من أسهم لثلاث شركات أمريكية ومقارنة الأداء مع خوارزميات خط الأساس خوارزميات الزخم (EG وFTRL) وخوارزمية التعلم المعزّز ARL تراوحت فترة التدريب من 2009 وحتى 2018 وفترة الاختبار هي سنة واحدة (2020). حيث أظهرت النتائج تفوق نموذج البحث MSPM على خوارزميات الزخم بمقدار مرة ونصف من حيث نسبة Sortino[1] للمحفظة الأولى وبمقدار 1.2 ضعف للمحفظة الثانية.

[1] تختلف عن نسبة شارب باستخدامها للانحراف المعياري للعوائد السالبة فقط مقارنة مع الانحراف المعياري لجميع العوائد المستخدم في نسبة شارب.

ما يميز هذه الدراسة

من أوجه الشبه بين الدراسات السابقة وهذه الدراسة هو استخدامها لخوارزميات إدارة المحفظة المستمرة في نطاق الاسـواق المالية. حيث تعد الدراسـة اسـتكمالاً للجهود السـابقة المبذولة من ناحية توظيفها لأهم ما وصلت له من أساليب ونتائج. وتتميز عنها من خلال ما يلي:

1) تطبيقها للتعلم المعزز العميق في سـوق دمشـق للأوراق المالية وبورصـة عمّان ومقارنة اداءه مع خوارزميات الزخم والمعاكسـة. حيث انه لم يتم تطبيق هذا النموذج واجراء مقارنة من هذا النوع في أي من الأسواق العربية بعد.
2) جمعها لعدد من معايير التقييم الأولية بهدف انتقاء الأسهم التي ستدخل ضمن المحفظة.

1-2- مشكلة البحث

تستخدم الخوارزميات الحاسوبية في مجال التداول الخوارزمي بهدف القيام بعملية بناء المحافظ الاستثمارية وإعادة هيكلتها دورياً عوضـا عن (أو بإشراف) المستثمر. حيث يوجد عدد كبير من هذه الخوارزميات التي تختلف فيما بينها من حيث الأداء وقابلية التطبيق. الأمر الذي يرفع من صـعوبة اختيار الخوارزمية الأمثل والتي تتناسـب مع تفضـيل المسـتثمر من عوائد وخطر من جهة، وقابلية تطبيقها في سـوق الأوراق المالية من جهة أخرى. بناء على ذلك يسعى البحث للإجابة عن السؤال التالي:

هل يوجد اختلاف في الأداء بين خوارزميات الإدارة المستمرة لمحافظ الاسهم؟

والذي يتفرع عنه الأسئلة التالية:

1) هل يختلف أداء المحـافظ التي يتم ادارتهـا باسـتخدام خوارزميـة التعلم المعزّز العميق عن أداء محافظ خوارزميات استراتيجية الزخم وفق نسبة شارب؟
2) هل يختلف أداء المحـافظ التي يتم ادارتهـا باسـتخدام خوارزميـة التعلم المعزّز العميق عن أداء محافظ خوارزميات استراتيجية المعاكسة وفق نسبة شارب؟
3) هل يختلف أداء المحافظ التي يتم ادارتها باسـتخدام خوارزميات اسـتراتيجية الزخم عن أداء محافظ خوارزميات استراتيجية المعاكسة وفق نسبة شارب؟

1-3- فرضيات البحث

للإجابة عن سؤال البحث سيتم اختبار الفرضية التالية:

يوجد اختلاف جوهري في الأداء بين خوارزميات الإدارة المستمرة لمحافظ الأسهم

وتتفرع عنها ثلاث فرضيات هي:

1) يوجد اختلاف في أداء المحافظ وفق نسبة شارب وذلك لصالح محفظة التعلم المعزّز العميق مقارنة مع محفظة استراتيجية الزخم.
2) يوجد اختلاف في أداء المحافظ وفق نسبة شارب وذلك لصالح محفظة التعلم المعزّز العميق مقارنة مع محفظة استراتيجية المعاكسة.
3) يوجد اختلاف في أداء المحافظ وفق نسبة شارب وذلك لصالح محافظ استراتيجية الزخم مقارنة مع محفظة استراتيجية المعاكسة.

1-4- أهمية البحث

تتبلور أهمية البحث في النقاط التالية:

1- يعد هذا البحث من الدراسات الأولى التي تستخدم التعلم المعزّز العميق في إدارة المحافظ في سوقي دمشق للأوراق المالية وبورصة عمّان.
2- مساعدة مستثمري سوق دمشق للأوراق المالية وبورصة عمّان على اختيار الخوارزمية الأمثل من خلال اختباره لعدة خوارزميات ومقارنة اداءه.
3- مساعدة نتائج المقارنة الباحثين في هذا المجال في التركيز على تطوير خوارزميات إدارة محافظ مستمرة خاصة بهذين السوقين بناء على أداء الخوارزمية الافضل.

1-5- أهداف البحث

يهدف هذا البحث إلى تطبيق خوارزميات التداول في كل من سوق دمشق للأوراق المالية وبورصة عمّان وتقييم أداء كل منها. ويمكن التوسع في ذلك كما يلي:

1- اختيار مجموعة من أسهم سوق دمشق وأسهم سوق عمّان للأوراق المالية التي ستدخل ضمن محافظ البحث من خلال استخدام عدد من معايير التقييم الأولي.

2- تطبيق نموذج تعلم معزز عميق إضافة إلى خوارزميات الزخم وخوارزميات المعاكسة لبناء محافظ استثمارية مستمرة مكونة من مجموعة الأسهم التي تم اختيارها.

3- مقارنة أداء الانواع الثلاثة من المحافظ من حيث معايير نسبة شارب Sharpe Ratio، القيمة التراكمية للمحفظة Accumulated Portfolio Value ونسبة الخسارة العظمى Maximum Drawdown Ratio.

1-6- منهج البحث

لتحقيق أهداف البحث وللوصل إلى أجوبة عن أسئلته واختبار فرضياته تم اتّباع المنهج الوصفي التحليلي وذلك من خلال ما يلي:

- دراسة خصائص المحافظ الاستثمارية والتعرف على التداول الخوارزمي ومفهوم إدارة المحافظ المستمرة إضافة إلى الخوارزميات المتعددة المبنية على استراتيجية الزخم واستراتيجية المعاكسة. كما يتضمن التعرف على مفهومي الشبكات العصبية الاصطناعية والتعلم المعزّز، بالإضافة إلى انواع، استخدامات، وطرق تدريب كل منهما ومن ثم خوارزمية التعلم المعزّز العميق المبنية على هذين المفهومين.
- تطبيق نموذج التعلم المعزّز العميق وخوارزميات الزخم والمعاكسة لبناء محافظ من أسهم سوق دمشق وسوق عمّان للأوراق المالية وذلك بالاستعانة بلغة البرمجة Python ومكتبة البرمجيات TensorFlow المتخصصة بتطبيقات التعلم الآلي ومكتبة Pandas التي تسهل عملية التعامل مع

البيانات وتحليلها. حيث تم تشـكيل عينة من أسـهم كل سـوق على حدا لبناء المحافظ الاسـتثمارية المستمرة ومقارنة نتائجها ومن ثم تحليل هذه النتائج.

1-7- مجتمع وعينة البحث

يتضمن مجتمع البحث أسهم سوق دمشق للأوراق المالية وأسهم بورصة عمّان، حيث تم اختيار عينة من أسهم كل سوق من خلال الاعتماد على عدد من معايير التقييم الأولية.

حيث تم تدريب نموذج التعلم المعزّز العميق باستخدام أسعار الأسهم اليومية للفترة بين 2011/01/09 و2018/10/02 وذلك لسوق دمشق، أما في بورصة عمّان فكانت فترة التدريب بين 2006/06/05 و 2017/10/05.

وكانت فترة الاختبار بين 2019/09/19 و 2020/07/30 لسوق دمشق وبين 2019/03/02 و 2020/07/30 يوماً لبورصة عمّان أي 315 و516 يوماً لكل سوق على التوالي. حيث تم اختيار هذه الفترات حسب البيانات المتوفرة لكل سوق مالي.

الفصل الأول – المحافظ الاستثمارية

مقدمة

يشير مصطلح الاستثمار إلى توظيف رؤوس الأموال للقيام بالمشاريع وتنشيطها والعودة بالمنفعة المالية على أصحاب الأموال والمنفعة الاقتصادية على المجتمع. تعد المحافظ الاستثمارية أحد اهم أدوات الاستثمار منذ ظهورها وحتى وقتنا الحالي، فقد ظهرت في الكثير من الأعمال البحثية وتمت تغطيتها من قبل باحثين في مجالات متعددة أهمها العلوم الاقتصادية والمالية.

ويشير مصطلح المحافظ الاستثمارية إلى مجموعة الأصول التي يتم ادارتها بشكل فردي مباشرة او عن طريق وسطاء ومحترفين تحت اسم "مدراء المحافظ"، حيث يشغل مدراء المحافظ مناصب في صناديق الاستثمار المشتركة، صناديق التحوط، البنوك، شركات التأمين، وغيرها من المؤسسات المالية.

يهدف هذا الفصل إلى التعريف بالمحافظ الاستثمارية وتوضيح أهميتها، وظائفها وأنواعها، ومن ثم سيتم شرح اساسيات تكوين المحافظ الاستثمارية. وبعد ذلك سيتم الانتقال إلى موضوع إدارة المحافظ المستمرة والتعرف على التداول الخوارزمي، التمثيل الرياضي لمشكلة إدارة المحافظ المستمرة، والخوارزميات المختلفة المستخدمة لمعالجة هذه المشكلة.

2-1- المبحث الأول - أساسيات المحافظ الاستثمارية

2-1-1- مفهوم المحافظ الاستثمارية

عرف (عبد الجواد والشـــديفات، 2006، ص.21) المحفظة الاســـتثمارية على أنها: "توليف من الأدوات الاستثمارية التي تضم أدوات مالية كالأسهم والسندات والمشتقات والودائع والقبولات والأذونات، كذلك تضم أدوات حقيقية كالعقارات والمعادن"

كما عرفها (العتيبي، 2005, ص.28) بأنها: "أداة استثمارية مركبة من الأصول الحقيقية والمالية، شريطة أن يكون هدف المستثمر تقليل مخاطر الاستثمار عن طريق تنويع الأصول المستثمر بها، وتنمية قيمتها السوقية"

وبالتالي يمكننا القول بان المحفظة الاســتثمارية هي كم من الأموال موزع بنســبة معينة على مجموعة من الأصول التي يمكن ان تكون أصول مالية او أصول حقيقية، من أهدافها تحقيق عائد للمستثمر من خلال الزيادة بالقيمة على مدى عمرها بالإضافة إلى تخفيض المخاطر التي يمكن ان يتحملها المستثمر في حال تخصيص كامل مبلغ الاستثمار لأصل واحد.

2-1-2- أهمية المحافظ الاستثمارية

تتجلى أهمية المحافظ الاستثمارية في النقاط التالية (قبنض، 2018، ص.75):

1) الأهمية المالية:

توفير الربح الملائم وتجنب إبقاء الأموال في حالة مجمدة، كما تتضمن الابتعاد عن مخاطر الاستثمار.

2) الأهمية الاقتصادية:

تساهم المحافظ الاستثمارية في تنشيط أسواق الاستثمار وتوفير البيئة والظروف الملائمة لجذب صغار المستثمرين (صناديق الاستثمار على سبيل المثال) يؤدي هذا الأمر إلى زيادة الأموال المستثمرة في الاقتصاد وتطوير القطاعات المختلفة، مما يساعد في تحقيق النمو الاقتصادي.

3) الأهمية الفردية:

تتمثل في توفير خبرات الإدارة الكفؤة ومهارات اختيار وتصفية الاستثمارات للمستثمر العادي الغير قادر أو من غير الملائم أن يقوم بإدارة الاستثمارات لوحده وذلك بسبب غياب المعرفة او ارتفاع تكاليف الاستثمار والتداول.

4) الأهمية على مستوى المنظمات:

تقوم المحافظ الاستثمارية بتوفير وسيلة لتوظيف الأموال الزائدة عن حاجة المنظمة لتجنب إبقائها على شكل أموال سائلة والتعرض لمخاطر انخفاض القيمة الزمنية للنقود. وبشكل معاكس توفر المحافظ التي تستثمر في أسواق النقد وسيلة للحصول على النقد السريع لتلبية احتياجات المنظمات من تسديد الديون ودفع المصاريف وغيرها.

5) الأهمية الاجتماعية:

تنعكس النقاط السابقة إيجابيا على المجتمع ككل عن طريق تحقيق النمو الاقتصادي الذي يؤدي إلى تحسين دخل الفرد والرفع من نوعية وكمية السلع والخدمات التي ستستطيع الشركات توفيرها لأفراد المجتمع.

2-1-3- وظائف المحافظ الاستثمارية

تتمثل اهم وظائف المحافظ الاستثمارية فيما يلي (خريوش، 1999، ص.20):

1) المحافظة على رأس المال الأصلي:

يمكن ان يواجه المستثمر خسائر رأس مالية ناتجة عن انخفاض القيمة السوقية للأوراق المالية، وذلك نتيجة المخاطر العالية التي قد تكون متأصلة في الورقة المالية. إلا أن هذه الأمر يجب ألا يصل إلى رأس المال الأصلي، حيث من الضروري أن يكون هنالك درجة مناسبة من التوازن بين العائد والخطر للمحافظة على رأس مال المحفظة.

2) استقرار تدفق الدخل:

يتطلب من مدير المحفظة المحافظة على تدفقات نقدية بنسبة معينة حسب نوع المحفظة، حيث يعد هذا الأمر مهما لبعض المستثمرين المساهمين في أصول المحفظة، تتكون هذه المصادر على سبيل المثال من توزيعات أرباح أسهم المحفظة وفوائد السندات التي تمتلكها المحفظة.

3) النمو في رأس المال:

أي من خلال زيادة قيمة أصول المحفظة عن طريق الاختيار المدروس والعقلاني للأصول ونسبتها ضمن المحفظة.

4) التنويع:

يعد التنويع من اهم مهام مدير المحفظة، حيث يترتب عليه اختيار الاستثمارات بشكل يولد أكبر قدر من العائد مع إبقاء الخطر المتولد عن مجموعة الاستثمارات في أقل درجة ممكنة.

5) قابلية تحويل أصول المحفظة إلى سيولة او نقدية:

حيث يجب على مدير المحفظة بناء محفظته بطريقة تمكنه من تحويل أصولها إلى نقدية بأقصى سرعة ممكنة وبأقل أعباء وخسائر عند بيع هذه الأصول. حيث يرتبط هذا الهدف مع هدف التنويع في أنه يجب التنويع في تواريخ استحقاق أصول المحفظة بحيث تكون السيولة متوفرة بشكل دائم.

تتراوح درجة أهمية هذه الأهداف حسب نوع المحفظة، إلا أن جميعها تهدف أولا إلى تحقيق التنويع بشكل أساسي، ومن ثم تقوم بالتركيز على أهداف ثانوية حسب الغاية من المحفظة واعتبارات الإدارة.

2-1-4- أنواع المحافظ الاستثمارية

يوجد العديد من المؤشرات التي تتيح تمييز المحافظ الاستثمارية، منها إمكانية المستثمرين المادية (مرونة الطلب)، اختلاف العائد المتوقع من إدارة المحفظة مقارنة مع أدوات بديلة، ظروف الاستثمار وما يتولد عنه من مخاطر عديدة، وكفاءة كل من إدارة الأسواق المالية وإدارة المحافظ المالية. حسب التصنيف العام للاستثمارات يمكننا التمييز بين أنواع محافظ الاستثمار التالية (هوشيار، 2009، ص.226):

1) محفظة الدخل:

يهدف هذا النوع من المحافظ إلى تحقيق عائد دوري وذلك من خلال التدفقات النقدية للأصل المالي بشكل أساسي إضافة إلى التغير بأسعار الأصول بشكل ثانوي. وبالتالي يكون التركيز فيها على الاستثمارات ذات الخطر المنخفض نسبياً، كالسندات الحكومية وأسهم وسندات الشركات الناضجة ذات التقييم الائتماني المرتفع. ويكون على مدير المحفظة المقايضة بين التحوط والمغامرة، لكن يفترض عدم الإفراط بالمغامرة لإبقاء العوائد ثابتة نسبياً. يفضل هذا النوع من المحافظ صغار المستثمرين والذين يعتمدون في معيشتهم على الدخل منها، أي المستثمرين الغير محبين للمخاطرة.

2) محفظة النمو:

تشمل أصول هذه المحافظ الأسهم التي تحقق نمو متواصل، حيث تكون معدلات النمو لمبيعات (وبالتالي إيرادات الشركات) المعيار الأساس لانتقاء أسهمها، حيث يفترض أن يكون معدل نمو لهذه الإيرادات أعلى من متوسط معدل نمو الشركات الأخرى. تحقق هذه المحافظ دخلها بشكل أساسي من الأرباح الرأسمالية الناتجة عن زيادة قيمة الأسهم.

3) المحفظة المتنوعة:

تقوم بتنويع محتوياتها لتشمل غالبية الأدوات الاستثمارية، وبالتالي تكون عوائدها متنوعة بين محافظ الدخل ومحافظ النمو، أي ان عوائدها تأتي من مصادر متعددة كالتوزيعات النقدية للأسهم وفوائد السندات على سبيل المثال، إضافة إلى النمو في القيمة السوقية لاستثماراتها.

4) المحفظة المتوازنة:

وهي محفظة تتكون من أوراق مالية تحقق التوازن بين الاستثمارات قصيرة الأجل (القابلة للتحويل السريع إلى سيولة ذات العوائد المنخفضة) والاستثمارات طويلة الأجل (صعبة التحويل إلى نقد بشكل سريع لكن عوائدها مرتفعة). وفي نفس الوقت تسعى هذه المحافظ إلى تحقيق التوازن بين العائد والمخاطرة.

5) المحافظ المتخصصة في القطاعات:

تستثمر هذه المحافظ في الأدوات المالية لشراكات قطاع معين كالطيران، التأمين، شركات الطاقة والنفط وغيرها. وهي غير شائعة الاستخدام أو مؤقتة لان التخصص في قطاع معين لا يتلاءم مع اهم وظيفة من وظائف المحافظ الاستثمارية وهي التنويع.

2-1-5- التنويع في المحافظ الاستثمارية

ان مبدأ التنويع يعد من الركائز الأساسية التي تقوم عليها المحافظ الاستثمارية، وهو عبارة عن دمج عدة أدوات مالية ضمن محفظة واحدة والهدف منه هو تجنب الاستثمار في أداة مالية واحدة الذي يؤدي إلى رفع المخاطر التي قد تواجهها المحفظة كالتقلبات قصيرة الأجل في سعر الأداة المالية. حيث يساعد التنويع في تخفيض أثر المخاطر التي تواجه الأداة الواحدة من خلال توزيعه على أنواع مختلفة من الأصول في قطاعات مختلفة.

تقسم المخاطر التي يمكن ان تتعرض لها المحفظة الاستثمارية إلى نوعين (Gallati, 2003, P.42):

1) مخاطر منتظمة Systematic Risk، وهي المخاطر التي لا يمكن تجنبها من خلال التنويع لأنها مخاطر تصيب السوق بأكمله أي أن جميع الأدوات المالية معرضة لها، وهي تتضمن خطر السوق، خطر أسعار الفائدة، خطر التضخم، وخطر أسعار الصرف.
2) مخاطر غير منتظمة Unsystematic Risk، وتسمى أيضا المخاطر المخصصة وهي المخاطر التي يمكن تخفيضها بالتنويع وتتضمن خطر الأعمال والأخطار المالية.

حيث يعد التنويع مهم جداً، ولكن يجب الحذر بعدم المغالاة فيه لأنه يمكن أن يتسبب بآثار عكسية منها (الهندي، 2003، ص.197):

1) صعوبة إدارة المحفظة:

تقوم عملية إدارة المحافظ المالية على التحليل المستمر لبيانات الأصول المالية التي تتكون منها المحفظة، وذلك للحصول على معلومات المنشآت العديدة التي تصدر هذه الأصول المالية وتقييم اداءها بهدف اتخاذ القرار بزيادة أو تخفيض حجم الاستثمار فيها. وبالتالي يؤدي زيادة عدد الأصول المختلفة ضمن المحفظة إلى ارتفاع درجة تعقيد القرارات التي يجب اتخاذها.

2) ارتفاع تكاليف البحث عن استثمارات جديدة:

إن زيادة عدد الأصول المختلفة ضمن المحفظة سيؤدي إلى رفع تكلفة تحليل بياناتها لأنها تتطلب وقت وتكلفة مرتفعة مقارنة مع المحافظ ذات التنويع المعتدل.

3) اتخاذ قرارات استثمارية غير سليمة:

يؤدي التويع الزائد إلى انخفاض فرصة إيجاد استثمارات تتناسب مع أصول المحفظة الحالية. حيث أن الاستثمار الجديد قد لا ينسجم مع الاستثمارات الحالية من حيث المخاطر وقد يتسبب بارتفاعها.

4) ارتفاع متوسط تكاليف الشراء:

واحدة من الاثار العكسية للتنويع أيضا هي ارتفاع عمولات ومصاريف التداول وبالتالي ستتأثر عوائد المحفظة بشكل سلبي.

إضافة إلى ذلك يؤدي ارتفاع التنويع في بعض المحافظ إلى اقترابها من محفظة السوق، أي انها لن تحقق أي عوائد إضافية عن السوق على الرغم من تكبدها لمصاريف تداول مرتفعة لشراء استثماراتها، حيث تتفوق عليها في هذه الحالة صناديق التداول التي تتبع مؤشر السوق.

2-1-6- أساسيات تكوين المحافظ الاستثمارية

يعتبر تخصيص الموجودات من الأمور ذات الأهمية العالية في عملية إدارة المحافظ الاستثمارية، حيث تمر هذه العملية بالمراحل الأربع التالية (ربابعة، 2009، ص.4):

1) وضع بيان السياسة الاستثمارية او خطة الاستثمار:

تعتبر الخطة الاستثمارية عامود البناء الأساسي الذي تشيَّد فوقه المحفظة الاستثمارية وتبنى عليه كافة القرارات، وتتضمن الخطة تحديد الأهداف الاستثمارية، درجة الخطر التي ستتحملها المحفظة والمحددات التي يجب وضعها في بداية طريق الاستثمار بشكل يتلاءم مع احتياجات وظروف المستثمر.

2) دراسة السوق والاتجاهات المستقبلية:

يتأثر أداء المحفظة بالكثير من العوامل المالية، السياسية والاقتصادية المختلفة التي هي خارجة عن سيطرة المستثمر لذلك فان عملية إدارة المحافظ الاستثمارية تتطلب متابعة مستمرة لهذه العوامل ودراستها للتنبؤ بها وتقييم أثرها على أداء المحفظة بهدف تخفيض الاثار السلبية لها بأكبر قدر ممكن.

3) تكوين المحفظة:

بالاعتماد على السياسة الاستثمارية ونتائج دراسة السوق يتم في هذه المرحلة تحديد كيفية تخصيص الأموال لتحقيق اهداف المستثمر، أي من خلال تحديد ماهي الأصول والأوراق المالية التي يجب شراؤها وما هي نسبة هذه الاستثمارات في المحفظة.

4) المتابعة وتقييم الأداء:

من المراحل الأكثر أهمية، حيث تتضمن مقارنة التوقعات والنتائج الفعلية والتعديل على الخطة والسياسة الاستثمارية إذا تطلب الامر ذلك.

2-2- المبحث الثاني - إدارة المحفظة المستمرة Online Portfolio Selection

تمهيد

في بداية الثمانينات بدأ التخصص فيما يدعى بالمالية الحاسوبية Computational Finance وهو فرع المعرفة الذي يقوم بالبحث وتطبيق النظريات المالية من خلال التقنيات الحاسوبية والأدوات التي طوّرت في مجال الإحصاء. واستمر هذا التطور وبشكل متسارع في السنوات الأخيرة.

يعد الاستثمار من أهم التخصصات الفرعية التي يتم دراستها ضمن المالية الحاسوبية، حيث يساعد الحاسوب في أتمتة العديد من المهام التي تساهم في الوصول إلى القرار الاستثماري. من خلال استخدام الأدوات الحاسوبية المتطورة، يستطيع المحلل المالي دراسة كم هائل من البيانات من أجل تحديد الأسهم التي يتم تداولها بأقل من قيمتها الذاتية واتخاذ قرار الشراء.

كما تتضمن المالية الحاسوبية ما يدعى بإدارة المحافظ المستمرة، وهو مجال البحث الذي يسعى إلى تعظيم المنفعة المادية من عملية توزيع رأس المال على مجموعة من الأصول. تتكون مجموعة الأصول التي يمكن الاختيار منها من أصلين على الأقل، وتمثل كامل الأصول المتاحة مفهوم السوق. تتسبب التغيرات في أسعار الأصول ضمن السوق في تغير مستوى خطر كل أصل، الامر الذي يتطلب من المستثمر اتخاذ القرار بتغيير نسبة هذا الأصل ضمن المحفظة.

وتعني كلمة "مستمرة" ضمن إدارة المحفظة المستمرة بأن عملية إعادة الهيكلة تحصل بشكل دوري ومؤتمت. وبشكل عام تلعب إدارة المحافظ المستمرة دورا هاماً في العديد من التطبيقات في مجال الاستثمار، منها إدارة الثروة المؤتمتة، إدارة صناديق التحوط والتداول الكمّي.

2-2-1- التحديات التي تواجه طرق الاستثمار التقليدية

في ظل التطورات العديدة الحاصلة على الصعيد الاقتصادي والتكنولوجي تواجه مؤسسات الاستثمار والمستثمرون الفرديون الذين يستخدمون طرق الاستثمار التقليدية عدد من التحديات التي يمكن التغلب عليها من خلال التداول الخوارزمي. من أهم هذه التحديات (Li and Hoi, 2018, P.4):

1) العدد الهائل من الأدوات المالية المتوفرة

وذلك من حيث الصنف والكم ضمن كل صنف. فمن جهة، تسببت الابتكارات في الهندسة المالية بتطوير أنواع جديدة من أدوات المالية كعقود مبادلة أسعار الفائدة، الخيارات، وعقود مبادلة خطر الإئتمان. ومن جهة أخرى، أدى تطور الاقتصاد العالمي إلى تضمين الأدوات المالية في العديد من الأسواق المالية الدولية. مما رفع من صعوبة عملية معالجة وتحليل هذه الاستثمارات على المستثمر الفردي.

2) عوامل التحيز في السلوك البشري

تعاني العديد من إستراتيجيات الاستثمار من التحيز الناتج عن الطبيعة الذاتية (غير الموضوعية) المتأصلة في الانسان. فعندما يتفاعل عاملي الخوف والطمع، يؤدي الأمر إلى اتخاذ قرارات غير مثالية. وفي الواقع، يقوم الكثير من المتداولين باستغلال هذه العوامل السلوكية. ولذلك من الأفضل للمستثمر الفردي او مؤسسة الاستثمار الفردية تخفيض هذه العوامل او الابتعاد عنها وعدم استغلالها.

3) التداول عالي التردد High-Frequency Trading

ظهر هذا التحدي نتيجة التطورات الحاصلة في مجال تكنلوجيا المعلومات، حيث أصبح بإمكان المتداول المستفيد من هذه التقنيات تنفيذ عمليات البيع والشراء في فترات تتراوح من ثواني إلى يومٍ واحد. وهو أمر أساسي عند التعامل مع البيانات التي تحصل خلال اليوم بما أنها أكبر حجما وأكثر سرعة من البيانات الأقل ترددا. فهي تتطلب أدوات ومنهجيات قادرة على الرد السريع، واتخاذ القرارات التي تتناسب مع التغيرات الحاصلة في السوق في الوقت المناسب لتجنب ضياع الفرصة.

في بعض الأحيان يستطيع المستثمر الفردي اكتشاف الفرص في هذا النوع من البيانات ولكنه في الغالب لن يكون قادراً على توليد أوامر البيع والشراء في الوقت المناسب لاستغلالها.

2-2-2- التداول الخوارزمي

يقوم التداول الخوارزمي على استخدام الخوارزميات الحاسوبية لبيع وشراء الأسهم، السندات، العملات، المشتقات وغيرها من الأدوات المالية. وهي قائمة على استراتيجية معينة تهدف إلى تحقيق هدف يتم وضعه من قبل المستثمر. يتم تطبيق الاستراتيجية من خلال ترجمتها إلى تعليمات رياضية محددة تتراوح من إرشادات بسيطة إلى أوامر عالية التعقيد، ليتم تطبيقها من قبل أجهزة الحاسوب (Kissel, 2013, P.1).

يتيح التداول الخوارزمي قابلية التفادي والاستفادة من التحديات السابقة، فالحاسوب قادر على التعامل مع كميات هائلة من البيانات مقارنة بالمستثمر العادي. باستخدام التداول الخوارزمي نستطيع تخطي عوامل التحيز الموجودة في السلوك البشري وبالتالي سنحصل دوماً على نفس النتيجة إذا قمنا بإدخال نفس المعطيات. وهو قادر على معالجة البيانات بسرعة هائلة وبالتالي يعد ملائم للاستفادة من البيانات عالية التردد واستخدامها لتحقيق الأهداف الاستثمارية (Li and Hoi, 2018, P.5).

يتفرع عن التداول الخوارزمي مجالين (Harris, 2003, P.33). المجال الأول هو المتعلق بطرف البيع ويهتم بالبحث في الخوارزميات التي تقوم بتقسيم أوامر البيع ذات الحجم الكبير على عدد من الاوامر صغيرة الحجم بشكل مؤتمت، وذلك بهدف تخفيض الأثر على سعر الأصل المباع في السوق إلى أقل قدر ممكن. تطبق المعرفة والأدوات التي تم تطويرها في هذا المجال من قبل السماسرة والوسطاء، من ضمنهم بنوك الاستثمار.

اما المجال الثاني فهو القائم على طرف الشراء، وهو مختص في إيجاد الخوارزميات الذكية القادرة على تحديد نوع الأصل، كميته، والوقت المناسب لشرائه، وبالتالي اتخاذ قرارات الاستثمار المناسبة لتحقيق

الهدف من الاستثمار. قد يكون الهدف تعظيم العوائد، تخفيض المخاطر، أو مزيج من الهدفين. يستخدم الخوارزميات في هذا المجال المستثمرون الأفراد، الشركات، والحكومات. وهو معني بمعالجة مشكلة إدارة المحفظة المستمرة بشكل أساسي.

2-2-3- التمثيل الرياضي لمشكلة إدارة المحفظة المستمرة

تتوقف إدارة المحافظ الاستثمارية المستمرة على عملية إعادة الهيكلة الدورية للأصول المالية في المحفظة، وذلك من خلال زيادة أوزان بعض الأصول وتخفيض أوزان أصول أخرى حسب توقعات مدير المحفظة.

في كل فترة زمنية t يمكن حساب قيمة المحفظة الاستثمارية من خلال المعادلة التالية (Jiang et al., 2017, P.5):

$$\boldsymbol{p_t = p_{t-1} r_t \cdot w_t}$$

$\boldsymbol{p_t}$: قيمة المحفظة في الفترة t

$\boldsymbol{p_{t-1}}$: قيمة المحفظة في الفترة $t-1$

$\boldsymbol{r_t}$: شعاع متَّجه (Vector) عوائد أصول المحفظة

$\boldsymbol{w_t}$: شعاع متَّجه أوزان الأصول ضمن المحفظة.

$\boldsymbol{r_t \cdot w_t}$: ناتج الجداء النقطي لشعاعي متجَّهي العوائد والأوزان ويمثل عائد المحفظة في الفترة t.

تخضع $\boldsymbol{w_t}$ للشرط (1):

$$\sum_{i=1}^{M} w_{t,i} = 1$$

حيث M هي عدد الأصول ضمن المحفظة.

وتحسب r_t من خلال المعادلة التالية:

$$\boldsymbol{r_t = \frac{v_t}{v_{t-1}}}$$

$\boldsymbol{v_t}$: شعاع متَّجه قيم (أسعار) أصول المحفظة في الفترة $\boldsymbol{t}$

وبالتالي يمكننا الحصول على قيمة المحفظة في نهاية فترة التداول من خلال (Jiang et al., 2017, P.6):

المعادلة 1

$$p_T = p_0 \prod_{t=1}^{T} r_t \cdot w_t$$

p_0 : يمثل قيمة الاستثمار الأولي أي قيمة المحفظة في الفترة $t = 0$

T : عدد فترات التداول

$\prod_{t=1}^{T} r_t \cdot w_t$: الجداء التسلسلي لعوائد أصول المحفظة المثقلة بأوزان هذه الأصول على مدى فترات التداول T، وهو معدل نمو المحفظة خلال عمرها.

إذا أردنا ادخال مصاريف التداول من عمولات وغيرها التي تدفع عند عمليات البيع والشراء للسماسرة، تصبح قيمة المحفظة كما يلي (المرجع السابق نفسه):

$$p_T = p_0 \prod_{t=1}^{T} \mu_t r_t \cdot w_t$$

μ_t : هي النسبة المئوية الباقية من قيمة المحفظة بعد إعادة الهيكلة، أي بعد خصم مصاريف التداول الناتجة عن عمليات بيع وشراء الأصول.

على سبيل المثال إذا كانت العوائد في الفترة t 800 وكانت قيمة مصاريف التداول في نفس الفترة 24 عندها تكون نسبة مصاريف التداول من إجمالي العوائد 3%، فتكون μ_t في هذه الحالة 97%.

الهدف الأساسي لمدير المحفظة هو تعظيم قيمة p_T. نظراً لكون عوائد أصول المحفظة r_t تابعة لتغيرات أسعار الأصول السوقية وهو أمر خارج عن سيطرة مدير المحفظة، الطريقة الوحيدة لتعظيم p_T هي اختيار w_t في كل فترة بحيث يكون وزن الأصول التي ستنخفض قيمتها أقل ما يمكن، ووزن الأصول التي سترتفع قيمتها أعلى ما يمكن. بهدف الوصول لهذه الأوزان يجب تقدير قيمة الأصول التي ستكون

عليها في المستقبل. أحد الطرق المتبعة في هذا التقدير هو استخدام خوارزميات التداول التي تقوم على استغلال معلومات متاحة عن الأصل المالي كالأسعار التاريخية بهدف التنبؤ بما سيكون عليه في المستقبل ومن ثم اتخاذ القرار بناء على نتيجة هذا التنبؤ.

2-2-4- خوارزميات إدارة المحفظة المستمرة

خلال السنوات السابقة، ومع التطورات التكنولوجية الحاصلة التي رفعت من أداء الحواسيب، قام العديد من الباحثين بالعمل على تطوير خوارزميات تقوم بعملية إدارة المحفظة المستمرة حيث يمكن تصنيف هذه الخوارزميات إلى خمسة أنواع (Li and Hoi, 2014, P.3) هي:

1) الخوارزميات المرجعية Benchmarks Algorithms:

وهي ابسط أنواع الخوارزميات. تعتبر كخط أساس Baseline وتستخدم بهدف تقييم أداء الخوارزميات الأخرى. تتضمن ما يلي:

a. استراتيجية الشراء والاحتفاظ Buy-and-Hold Strategy، والتي تقوم على استثمار المبلغ الأولي في تركيبة معينة من الأصول وإبقاء هذه التركيبة طوال عمر المحفظة دون تغيير أوزان الأصول.
b. استراتيجية تداول السهم الأفضل Best Stock Strategy، وهي شبيهة باستراتيجية الشراء الطويل، إلا أنه لا يتم التنويع فيها، بل يتم استثمار كامل المبلغ في الأصل ذو الأداء الأفضل وابقاءه طوال عمر المحفظة.
c. استراتيجية محافظ إعادة التوازن الثابت[2] Constant Rebalanced Portfolios والتي تعتبر أكثر تعقيدا من الاستراتيجيتين السابقتين. وهي تقوم على تحديد تركيبة معينة من الأصول في

[2] إعادة التوازن هنا تعني أنه خلال عمر المحفظة ستتغير أوزان الأصول ضمنها بسبب التغيرات بأسعارها في السوق، وبالتالي يكون من الضروري إعادة شراء وبيع بعض الأصول للإبقاء والمحافظة على الاوزان الأساسية.

بداية عمر المحفظة والقيام بالإبقاء على نفس الأوزان حتى نهاية فترة الاســـتثمار. حيث تؤدي عوامل السـوق إلى التغيير من وزن كل أصـل ضـمن المحفظة، فتقوم الخوارزمية بحسـاب كمية الأصول التي يجب بيعها وتلك التي يجب شراؤها وتنفيذ العمليات من أجل إبقاء التوزيعات ذاتها.

2) خوارزميات استراتيجية الزخم Momentum Strategy Algorithms

وتدعى أيضـا خوارزميات اتبع الرابح. تقوم اسـتراتيجيات الزخم على اسـتغلال الاتجاه السـابق في الأسـواق من خلال الاستثمار في الأصول التي حققت أفضل أداء، حيث يتأمل مستثمر الزخم أن يدخل السوق في الوقت الذي يبدأ فيه الأصل الارتفاع في القيمة.

3) خوارزميات استراتيجية المعاكسة Contrarian Strategy Algorithms

كما تدعى باســـتراتيجية اتبع الخاســـر، اســـتراتيجية عكس المتوســـط او اســـتراتيجية التخالف. تأخذ هذه الاستراتيجية اتجاه السوق المعاكس، أي أن مستثمر المعاكسة يبحث عن الأصول ذات الاداء الأسوأ على افتراض أن هذا الأداء غير واقعي وأن اتجاهها سينعكس في المستقبل.

4) خوارزميات مطابقة النمط Pattern Matching Algorithms

تتكون هذه الخوارزميات عادة من خطوتين، الخطوة الأولى تدعى اختيار العينة وهي تعتمد على اكتشـــاف الانماط ضـمن سـلسـلة الأسـعار التاريخية للأصـل المالي. وفي الخطوة الثانية تقوم الخوارزمية باسـتغلال الانماط التاريخية ومحاولة عكسها على الفترات الحالية بهدف تعظيم عائد المحفظة المالية.

5) خوارزميات التعلم الفوقي Meta-Learning

تدمج بين إثنين أو أكثر من أنواع الخوارزميات السـابقة، كما تتضـمن خوارزميات التعلم الآلي. كما يمكن تأخذ مخرجات خوارزميات تعلم آلي أخرى كمدخلات لها وتقوم بالتعلم منها.

في هذا البحث سيتم استخدام ثلاثة من خوارزميات استراتيجية الزخم وثلاثة خوارزميات مناقضة. إضافة إلى واحدة من خوارزميات التعلم الآلي (الفصل الثاني) بهدف التجريب العملي على سوقي دمشق وعمّان ومقارنة النتائج (الفصل الثالث).

2-2-5- خوارزميات استراتيجية الزخم

1) المحافظ الشاملة Universal Portfolios

تعد خوارزمية UP او المحافظ الشاملة (Cover, 1991, P.1-29) شبيهة باستراتيجية الشراء والاحتفاظ Buy and Hold، حيث تحتفظ بعدد من الأصول على مدى عمر المحفظة. إلا أنها تختلف عنها بكون هذه الأصول محافظ بحد ذاتها، وبذلك تكون شبيهة بصناديق الاستثمار التي تستثمر بصناديق استثمار أخرى. المحافظ التابعة لهذه الخوارزمية هي "محافظ إعادة التوازن الثابت" Constant Rebalanced Portfolios والتي تقوم على اختيار أوزان معينة لأصولها والحفاظ على هذه الأوزان حتى نهاية عمر المحفظة. حيث يمكن بداية إعطاء أوزان عشوائية لأصول هذه المحافظ.

تبدأ المحفظة الاساسية باتباع تنويع ساذج من خلال توزيع رأس المال بشكل متساوي على جميع محافظ إعادة التوازن. في كل فترة إعادة هيكلة لاحقة تحسب الأوزان الجديدة في المحفظة الأساسية تبعا لمدى نجاح المحافظ التابعة. وذلك من خلال زيادة الأوزان في المحافظ ذات راس المال المرتفع، وتخفيضها في المحافظ التي لم تحقق نمو مشابه.

إذا افترضنا أن المحفظة x تحتوي على الأصل a بنسبه عالية، والمحفظة y تحتوي على نفس الأصل a ولكن بنسبة منخفضة. فإذا حقق الأصل a عوائد مرتفعة، ستنمو المحفظة x بمعدل أعلى من المحفظة y بافتراض بقاء بقية الأصول في كل من المحفظتين على حالها. وبالتالي عند إعادة الهيكلة ستقوم المحفظة الأساسية بزيادة وزن المحفظة x وتخفيض وزن المحفظة y، متبعة بذلك استراتيجية الزخم.

2) الميل الأُسي Exponential Gradient

خوارزمية الميل الأُسي (Helmbold, 1998, P.325) تقوم على تحديث أوزان أصول المحفظة للفترة التالية باستخدام أوزان الفترة الحالية، بحيث أنه إذا كان معدل النمو لإحدى أصول المحفظة أكبر من وسطي معدل النمو للمحفظة الكاملة في الفترة الحالية، يجب زيادة نسبة هذا الأصل في المحفظة. حيث يستخدم مقياس معدل التعلم في هذه الخوارزمية لزيادة شدة التعديل الحاصل في فترات التداول.

تحوي خوارزمية الميل الأُسي على حد يضاف إلى معادلات تحديث الأوزان يدعى بحد التسوية. يستخدم في هذه المعادلات حد الأنتروبيا النسبية Relative Entropy كحد تسوية، ويهدف إلى تخفيض الفرق بين أوزان الفترة الحالية والفترة التالية إلى أقل ما يمكن.

تحاول هذه الخوارزمية –بشكل مشابه للمحافظ الشاملة– الوصول إلى نتائج قريبة من محفظة إعادة التوازن المستمر الأفضل BCRB، وهي محفظة نظرية تبنى من البيانات الفعلية عن فترة سابقة بحيث تحقق أفضل العوائد في تلك الفترة. يدعى المؤشر الذي يقارن بين معدل النمو اللوغاريتمي لكل من محفظة BCRB وأي خوارزمية أخرى بحد الندم Regret Bound.

تحقق خوارزمية الميل الأُسي حد ندم أسوأ قليلا من حد ندم خوارزمية المحافظ الشاملة لكنها تتفوق عليها بكونها أسرع بالتنفيذ وبالأخص عندما يكون عدد الأصول والفترات كبيراً.

3) عتبة نيوتن المتصلة Online newton step

هي شبيهة بخوارزمية الحد الأُسي إلا أنها تستغل المشتق الثاني من معادلات تحديث الأوزان. وتحقق حد ندم أسوأ بقليل من المحافظ الشاملة لكنها أسرع بالتنفيذ مقارنة بالمحافظ الشاملة والميل الأُسي (Agarwal, 2006).

2-2-6- خوارزميات استراتيجية المعاكسة

1) خوارزمية انعكاس المتوسط العدوانية Passive Aggressive Mean Reversion Algorithm

تستغل خوارزمية PAMAR (Dochow, 2015, P.97) ارتداد متوسط أسعار الأصول من خلال إتّباع مبدأ معاقبة الأصول التي حققت عوائد أعلى من متوسط عائد السوق في الفترة السابقة، وذلك عن طريق تخفيض نسبتها في المحفظة.

تبدأ المحفظة بأوزان متساوية لجميع الأصول، وتعدل أوزان الأصول في الفترات اللاحقة من خلال الوزن الحالي إضافة إلى فرق عائد الأصل عن متوسط عائد السوق مضروباً بدالة مضاعف. تحسب دالة المضاعف من خلال مقياس تقلب عوائد السوق إضافة إلى دالة خسارة الأصل والهدف منها هو رفع أو تخفيض حجم التعديل.

يوجد عدة أنواع لدالة المضاعف، بعضها يتيح للخوارزمية استخدام مؤشر للتغيير بين وضعين، هما الوضع العدواني والوضع غير العدواني. يختلف الوضعان عن بعضهما من خلال شدة درجة معاقبة نجاح الأصل، حيث أن تخفيض قيمة هذا المؤشر سيؤدي إلى زيادة عدوانية الخوارزمية وبالتالي تخفيض وزن الأصل الناجح في الفترة التالية بنسبة كبيرة في المحفظة، والعكس صحيح للوضع غير العدواني الذي تكون قيمة تخفيض الوزن فيه أقل.

2) خوارزمية انعكاس المتوسط للوسطي المتحرك المتصل Online Moving Average Mean Reversion Algorithm

خوارزمية OLMAR (المرجع السابق نفسه، ص.101) اختصاراً، هي توسّع لخوارزمية PAMAR التي تأخذ بعين الاعتبار فترة واحدة فقط من فترات التداول اللاحقة، بينما تقوم OLMAR بإدخال أكثر من فترة في المعادلة عن طريق حساب المتوسط المتحرك لأسعار الأصول ضمن نافذة زمنية محددة.

3) خوارزمية انعكاس الوسيط المتينة Robust Median Reversion Algorithm

تعد خوارزمية RMR (Dochow, 2015, P.102) تطويراً لخوارزمية OLMAR، لكن بدلاً من المتوسط المتحرك لأسعار الأصول كمؤشر للتنبؤ بقيمتها في فترة التداول التالية واختيار نسبتها في المحفظة، تستخدم خوارزمية RMR الوسيط الاحصائي لأسعار الأصول على شكل شعاع متّجه للقيام بذلك.

تأخذ الخوارزمية كمدخلات لها، حجم نافذة التداول، معامل انعكاس المتوسط، قيمة عتبة الانعكاس ومعامل تكرار، حيث تقدر قيمة متجهات أسعار الأصول على عدة دورات يحدد عددها بمعامل التكرار، وذلك باستخدام المعلومات المتاحة ضمن نافذة التداول.

الفصل الثاني – التعلم الآلي

مقدمة

في النهج التقليدي للبرمجة يتم إخبار الحاســـوب ما عليه فعله من خلال تحليل المشـــاكل إلى عدّة مهام بسيطة ومعرّفة بشكل واضح ومفصّل ومن ثم ترجمتها إلى تعليمات برمجية ليستطيع الحاسوب تنفيذها. يوجد العديد من المشاكل المعقّدة والغير قابلة للتبسيط، كالتنبؤ بسلوك المستهلك أو التصنيف والتصرف مع الرسائل المزعجة، والتي لا يمكن حلّها بشكل مباشر باستخدام الطرق التقليدية، لكن يمكننا التوصل إلى حل تقريبي لها من خلال البيانات الفعلية المتوفرة عنها. حيث نســـتطيع تدريب الحاســـوب على هذه البيانات لنمذجة الظاهرة والوصـــول إلى حل للمشـــكلة. قد تكون البيانات المتاحة غير كافية لتغطية جميع أوجه المشـــكلة، ومن المحتمل أن جمع بيانات عن كامل أوجه المشـــكلة هو أمر مكلف أو غير ممكن إطلاقاً. لكن في كثير من الحالات، يعد وجود نموذج تقريبي أمر كافي لحل المشـــكلة. وهذا هو جوهر التعلم الآلي.

التعلم الآلي هو المجال التطبيقي للذكاء الاصـــطناعي المعني بتطوير أنظمة قادرة على محاكاة الذكاء البشـــري. عندما نعطي صـــفة الذكاء لنظام معين، نعني أن هذا النظام قادر على التغير للتأقلم مع بيئة ديناميكية، أي انه قادر على التعلم منها للتعرف على الأنماط واتخاذ القرارات بأقل تدخل ممكن من مصمم النظام.

في هذا الفصل سيتم التعرف على نشأة، آلية عمل، وطريقة تدريب كل من الشبكات العصبية الاصطناعية والتعلم المعزّز. ومن ثم التطرق إلى التعلم المعزّز العميق الذي يقوم بدمج هاتين التقنيتين ومعرفة آلية تطبيقه في مجال إدارة المحافظ المستمرة.

3-1- المبحث الأول - الشبكات العصبية الاصطناعية Artificial Neural Networks

تمهيد

الشبكات العصبية الاصطناعية هي نماذج أو أنظمة تعكس سلوك الدماغ البشري، تستخدم في التصنيف وحل المشاكل الشائعة في مجال الذكاء الاصطناعي. حيث يتم عرض بيانات فعلية على الشبكة تتضمن المدخلات والمخرجات، ويكون عليها تعلم كيفية ربطها والتوصل إلى حل للمشكلة.

على سبيل المثال أحد هذه المشاكل هي تمييز الوجوه لنظام أمان إلى نوعين من الأشخاص (أشخاص مخوّلين وأشخاص غير مخوّلين للدخول) أو مشكلة التعرف على الأرقام من خلال الصور، كل من هاتين المشكلتين تتضمن التنبؤ بتصنيف معين، تدعى نتيجة التنبؤ بمخرجات الشبكة (إما مخوّل أو غير مخوّل للمشكلة الاولى، وهي الرقم المتعرف عليه في المشكلة الثانية)، يتم ذلك من خلال عرض آلاف الصور على الشبكة مع المخرجات الصحيحة المقابلة لكل من هذه الصور، بحيث تستطيع الشبكة لاحقاً تصنيف أي صورة شبيهة بها حتى لو لم تكن هذه الصور نفسها قد مرّت عليها سابقاً. هنالك ميزتين أساسيتين لحل هذه المشاكل بهذه الطريقة عوضاً عن برمجة الحل بالشكل التقليدي. أولاً توفير مقدار كبير من وقت وجهد لتصميم الحل، وثانياً، إمكانية استخدامها مع المشاكل التي لا يمكن تحديدها بشكل دقيق مسبقاً (Anthony, 1999, P.1).

وبالتالي الشبكات العصبية هي الأكثر ملائمة لحل المشاكل المعقدة، صعبة التحديد، غير الخطيّة، والتي تتضمن الكثير من المتغيرات. هذه المشاكل موجودة بوفر في مجالات الطب، العلوم المالية، الأمن الإلكتروني وغيرها، أي هي مشاكل ذات أهمية وشأن أساسيين.

3-1-1- تعريف الشبكات العصبية الاصطناعية

عرف (Haykin, 1999, P.2) الشبكة العصبية بأنها: "هي معالج يتكون من وحدات معالجة بسيطة لها ميل طبيعي لتخزين المعرفة التجريبية وإتاحتها للاستخدام، وهي تشابه الدماغ من ناحيتين، الأولى أن الشبكة تحصل على المعرفة من البيئة المحيطة من خلال عملية التعلم، والثانية أنه يتم استخدام الوصلات التي تربط بين العصبونات، والتي تعرف بالأوزان، لتخزين هذه المعرفة"

كما عرفها (Mehlig, 2021, P.1) بأنها: "بنى مستوحاة من ديناميكيات شبكات الخلايا العصبية في الدماغ. تستخدم خوارزميات الشبكات العصبية نماذج عصبية مبسطة ومثالية. ... وهي جزء من التعلم الموجّه الذي يتم تدريب الشبكات فيه على ربط عدد من المدخلات مع المسميات (المخرجات) الخاصة بها"

وبالتالي يمكننا القول بأن الشبكات العصبية هي إحدى أشكال التعلم الموجّه ذات بنية مستوحاة من الجهاز العصبي عند الإنسان بحيث تحاكي بعملها العصبونات العضوية من خلال إرسالها للمعلومات ضمن شبكة من هذه العصبونات.

3-1-2- نشأة وتطور الشبكات العصبية الاصطناعية

يعود تاريخ الشبكات العصبية لِـ McCulloch-Pitts اللذان وضعا النموذج المبدئي في عام 1943. شكّل مزيج الخوارزميات والرياضيات الذي تم الاعتماد عليه ما يعرف بمنطق العتبة Threshold logic، الذي أُستخدم في نموذجهم (MCP) لمحاكاة عملية التفكير، إلا انها لم تكن قادرة على التعلم (Géron, 2019, P.278). فيما بعد جاءت النظرية الهيبية Hebbian Theory والتي تعد محاولة لتفسير الليونة العصبية، حيث تقوم على فكرة أنه إذا قامت الخلية العصبية X بتفعيل الخلية العصبية Y المتصلة بها، ستزيد شدة الاتصال بين الخليتين وستصبح الخلية X قادرة على تفعيل Y بسهولة في فترات لاحقة (Liston et al., 2011, P.1170).

في ذلك الوقت كان عالم النفس Frank Rosenblatt يعمل على دراسة نظم اتخاذ القرارات البسيطة، وتوصل في مرحلة ما إلى فكرة البيرسيبترون Perceptron في عام 1958، وهي عبارة عن نموذج خلية عصبية مبني على منطق العتبة المطور من قبل McCulloch–Pitts لكنها قادرة على تعلم الأوزان الياً من خلال مقارنة المخرجات المحسوبة مع المخرجات المتوقعة خلال عدة دورات، وتعديل الأوزان بشكل تدريجي. حيث يعد البيرسيبترون –وما أصبح لاحقا بعد التطويرات التي حصلت عليه– حجر البناء الرئيسي لجميع الشبكات العصبية.

في عام 1959 نجح كل من Bernard Widrow و Marcian Hoff بتطوير أول شبكة عصبية اصطناعية قادرة على حل مشكلة واقعية (Widrow and Hoff, 1960, P.96) تدعى هذه الشبكات Adaline، وتستخدم نوع من الخلايا العصبية شبيه بالبيرسيبترون إلا أنه مختلف من حيث المخرجات.

ساهم هذا النجاح برفع مستوى الحماس ضمن الوسط العلمي اتجاه امكانية الشبكات العصبية في حل نمط جديد من المشاكل. استمر هذا الوضع حتى 1969 حيث نشر Marvin Minsky كتابه "Perceptrons" الذي وضح فيه أن البيرسيبترون المطور من قبل Rosenblatt غير قابل للاستخدام بشكل فعال في شبكات عصبية من عدة طبقات، حيث أن عملية الوصول إلى قيمة الأوزان الموزعة على عدة طبقات من خلال مخرجات الشبكة الموجودة في آخر طبقة يمكن أن يأخذ عدد لا نهائي من التكرارات (Minsky, 1969).

أدى عمل Minsky بتوضيحه لهذه المشكلة إضافة إلى عدد من المشاكل الأخرى التي عانت منها الشبكات العصبية في ذلك الوقت إلى اقناع عدد كبير من الباحثين ومؤسسات التمويل البحثي بأن الاستمرار في البحث في هذا المجال لن يعود باي فائدة، مما أدى إلى توقف تمويل هذه الأبحاث ودخلت الشبكات العصبية بفترة جفاف تدعى شتاء الذكاء الاصطناعي AI Winter (Umbrello, 2021, P.8).

في نهاية هذه الفترة تم الوصـول إلى أحد أهم الإنجازات في هذا المجال وهو خوارزمية الانتشـار العكسـي Backpropagation والتي كانت موجودة كفكرة بين الباحثين خلال السـتينات إلا أنه لم يتم اعتمادها وتطويرها بشـكل رسـمي حتى عام 1974 (Werbos, 1994) ، وتعتبر –بالإضـافة إلى خوارزمية Gradient descent– العمود الفقري للشبكات العصبية الاصطناعية.

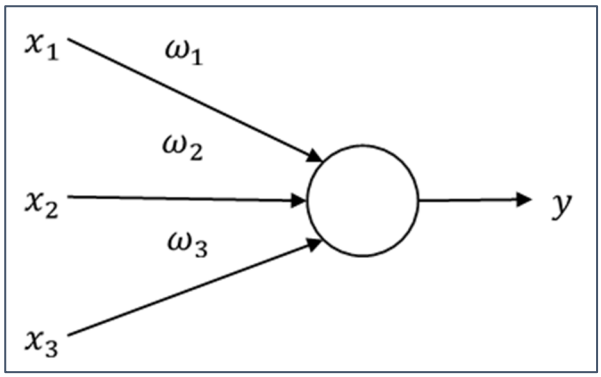
وفي عام 1975 تم ابتكار شـبكة Cognitron (Fukushima, 1975, P.121) والتي صـممت خصيصا للتعرف على الأنماط، طورت هذه الشبكة لاحقاً لتصبح Neocognitron والتي الهمت الشبكات العصـبية التلافيفية Convolutional Neural Networks (CNN) التي تعد أحد اهم أنواع الشـبكات العصـبية حيث تسـتخدم حالياً بشـكل واسـع في مجال تصـنيف الصـور. كما تم الوصـول إلى الشـبكات العصـبية التكرارية Recurrent Neural Networks (RNN) في 1986 والتي تعد من الأنواع المهمة أيضا ولها استخداماتها في مجال التعرف على الصوت.

تعد كل من شبكات RNN، CNN وغيرها كالـ DBN والـ LSTM المستوحاة من RNN تابعة للتعلم العميق Deep Learning وهو تصـنيف فرعي ضـمن الشـبكات العصـبية الاصـطناعية، تتميز عن الشـبكات العادية باحتوائها على طبقتين أو أكثر من الطبقات المخفية، الأمر الذي يرفع من فعاليتها في حل مشـاكل أكثر تعقيداً على حسـاب زيادة صـعوبة تدريبها. أي انها تتطلب قوة حاسـوبية أكبر من أجل تقدير معلمات الشـبكة. إلا أن هذا الأمر أصـبح ممكنا في بداية القرن الجديد مع التطورات الحاصـلة في الأجهزة الحاسوبية.

3-1-3- بنية والية عمل عصبون البيرسيبترون

من أقدم أنواع العصبونات هو البيرسيبترون الذي طوره Frank Rosenblatt، يعتبر البيرسيبترون وحدة اتخاذ قرار بسيطة، حيث يأخذ عدد من المدخلات x_n التي يجب أن تكون متغيرات منطقية حصراً (أي تأخذ قيمة واحدة من قيمتين, 1 او 0)، يقوم بمعالجة هذه المدخلات ويعطي قيمة 1 او 0 كمخرجات.

الشكل 1 - عصبون البيرسيبترون



المصدر: من إعداد الباحث

لحساب قيمة المخرجات اقترح Rosenblatt فكرة الأوزان w_n، والتي هي أرقام حقيقية تعبر عن شدة الرابط بين كل واحدة من المدخلات والقيمة المخرجة، حيث يتم حساب المجموع المرجح للمدخلات من خلال المعادلة:

$$\sum_{i=1}^{n} x_i w_i$$

بعد حساب هذا المجموع، يتم مقارنة الناتج مع قيمة عتبة Threshold فاذا كانت النتيجة أكبر منها يعطي البيرسيبترون قيمة 1، ويعطي قيمة 0 في حال كانت النتيجة أصغر او تساوي العتبة. فعلى سبيل المثال، بافتراض أن أحد المستمرين يريد استثمار 100 وحدة نقدية في شركة معينة، ولاتخاذ القرار هنالك عدد من العوامل التي يجب النظر فيها:

x_1 : هل كانت نسبة ربح السهم في السنة الماضية أكبر من 10%؟ (نعم = 1، لا = 0)

x_2 : هل سعر السهم أقل ام أكثر من 5 وحدة نقدية؟ (أقل = 1، أكثر = 0)

x_3 : هل تتبع الشركة سياسة خضراء باتخاذها لقرارات صديقة للبيئة؟ (تتبع = 1، لا تتبع = 0)

ولنفترض أن العامل الاول ذات أهمية متوسطة أي أنه يفضل أن تكون نسبة ربح السهم أكثر من 10%، وأنه يريد التنويع في استثماراته، لذا يريد سعر السهم أن يكون أقل ما يمكن ليستطيع شراء عدة أسهم، وأنه يفضل أن تكون سياسات الشركة خضراء، إلا أن هذا العامل ليس بنفس درجة الأهمية كالعوامل الأخرى. لذلك أعطيت الأوزان لكل من هذه العوامل على الشكل التالي:

$$\omega_1 = 4,\ \omega_2 = 8,\ \omega_3 = 2$$

وبالتالي لكي يتم اتخاذ القرار، يتم أولاً احتساب مجموع هذه العوامل المرجح بالأوزان. وبعد التحقق من بيانات الشركة كانت العوامل على الشكل التالي:

$$x_1 = 1,\ x_2 = 0,\ x_3 = 1$$

فيكون المجموع المرجح: $1 \times 4 + 0 \times 8 + 1 \times 2 =$ **6**

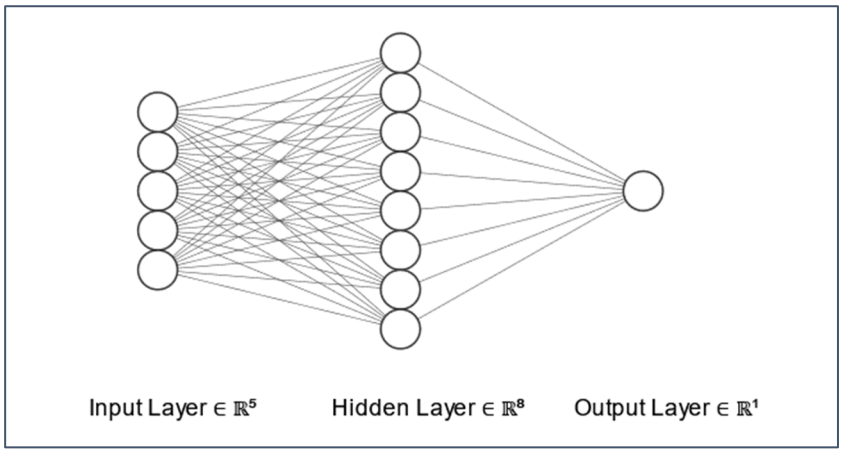
فاذا افترضنا أن العتبة هي 5 على سبيل المثال، يتم مقارنة مجموع العوامل المرجح بالأوزان مع هذه العتبة، فاذا كان المجموع أكبر من العتبة يكون القرار 1، أي سيقوم المستثمر بشراء هذا السهم.

من الواضح أن البيرسيبترون لا يشكل نموذج كامل لعملية اتخاذ القرار عند البشر، ولكن يوضح المثال قدرته على اخذه لعدة عوامل بعين الاعتبار والتفضيل بينها من أجل اتخاذ القرار. وبالتالي من المعقول الافتراض أنه إذا قمنا بجمع أكثر من عصبون على شكل شبكة سنستطيع رفع درجة تعقيد القرارات والاقتراب من درجة تعقيدها عند البشر.

3-1-4- بنية الشبكات العصبية الاصطناعية

تتكون الشبكة العصبية الاصطناعية من عدة عصبونات، يمكن أن تكون من نوع البيرسيبترون، أو أنواع أخرى كما سنرى في الفقرة التالية. تجمع هذه العصبونات ضمن طبقات. تتألف أبسط أنواع الشبكات العصبية من ثلاث طبقات (Nielsen, 2015, P.11) كما هو موضح في الشكل (2):

الشكل 2 - بنية الشبكة العصبية



المصدر: من إعداد الباحث

1) طبقة المدخلات: تدعى العصبونات في هذه الطبقة عصبونات الإدخال حيث تتلقى البيانات الخام.
2) الطبقة المخفية: هي جميع الطبقات التي تقع بين طبقة المدخلات وطبقة المخرجات أي أنه يمكن أن يكون هنالك أكثر من طبقة مخفية ضمن الشبكة العصبية.
3) طبقة المخرجات: هي نتيجة أو حل الشبكة للمشكلة، يمكن أن تكون تنبؤ أو تصنيف معين لمجموعة بيانات.

يكون كل عصبون في كل طبقة متصل مع جميع العصبونات من الطبقة السابقة (عدى عصبونات طبقة المدخلات) لتشكل بذلك شبكة من الوصلات تدعى أوزان. ويُحَدَد كل من عدد العصبونات في كل طبقة، عدد الطبقات المخفية، ونوع الطبقات المخفية حسب طبيعة المشكلة ونوع البيانات المتوفرة.

3-1-5- تدريب الشبكات العصبية الاصطناعية

الهدف الأساسي من تدريب الشبكات العصبية هو الوصول إلى أوزان الشبكة التي تجعلها قادرة على إعطاء المخرجات الصحيحة. لتكون الشبكة قادرة على التعلم، كل تغير بسيط في الأوزان يجب أن ينتج عنه تغير بسيط في المخرجات، لكن هذا الأمر غير ممكن في شبكة مكونة من عصبونات البيرسيبترون. حيث أن أي تغير بسيط في أوزان شبكة البيرسيبترون يمكن ان يؤدي إلى قلب كامل مخرجات البيرسيبترون (من 0 إلى 1 أو بالعكس). كما سيؤدي هذا التغيير إلى التأثير في سلوك الشبكة بطريقة معقدة، على سبيل المثال، إذا كانت الشبكة تستعمل في تصنيف الصور، وقمنا بتغيير الأوزان بطريقة معينة، يمكن أن يؤدي هذا التغيير إلى تصنيفها لإحدى الصور بطريقة صحيحة لكن من الممكن أن يجعلها غير قادرة على تصنيف بقية الصور بشكل صحيح.

لتفادي هذه المشكلة يمكن استخدام عصبون السيغمويد Sigmoid حيث يشابه هذا العصبون البيرسيبترون من حيث الشكل إلا أن مدخلاته يمكن أن تأخذ أي رقم حقيقي بين 0 و1. ويتم حساب مخرجاته بنفس الطريقة (المجموع المرجح للأوزان والمدخلات) إلا أنها تمر بدالة تدعى دالة سيغمويد Sigmoid، حيث تكون مخرجاتها أيضا رقم حقيقي بين 0 و1. وبذلك نستطيع الاعتماد على المبدأ السابق في تدريب الشبكات (Nielsen, 2015, P.7).

تقسم الطرق المستخدمة للتدريب في مجال التعلم الآلي إلى ثلاث مناهج تعلم مصنفة حسب الإشارة (نوعية التغذية الراجعة) التي يتعلم منها للنموذج (Sutton and Barto, 2018, P.2) إلى:

1) التعلم الموجّه Supervised Learning:

هو التعلم من مجموعة أمثلة مصـــنفة (بيانات)، كل مثال هو عبارة عن وصـــف لحالة معينة (مدخلات) مرفق مع تصـــنيف لهذه الحالة أو التصـــرف الصـــحيح الذي يجب أن يتخذه النموذج في هذه الحالة (مخرجات). الهدف من هذا المنهج هو تدريب النموذج على تعميم المعرفة التي سـيحصـل عليها من هذه البيانات إلى جميع الحالات المشابهة الغير موجودة في البيانات.

اهم الطرق المسـتخدمة في التعلم الموجّه هي نماذج الانحدار بأنواعها، Support-Vector Machines والشبكات العصبية الاصطناعية.

2) التعلم غير الموجه Unsupervised Learning:

يقوم على إيجاد بنى مخفية في بيانات غير مصنفة. في هذا النوع من التعلم يعطى النموذج مدخلات فقط دون المخرجات الفعلية، ويكون عليه إيجاد ملامح معينة في البيانات تسمح له إما بتجميعها إلى مجموعات معينة حسب هذه الملامح، بربطها بتصنيف ما، أو بتبسيطها من خلال ما يعرف بتخفيض الأبعاد. بعض هذه النماذج هي k-Means، Hierarchical Clustering و Local Outlier Factor.

3) التعلم المعزّز (او التعلم بإعادة التعزيز) Reinforcement Learning:

وهو إطار يشمل عدد من المكونات كعميل او متحكم يقوم على تعظيم دالة مكافأة، ويستفيد من معلومات البيئة لاتخاذ قراراته. لا يتم إعطاء العميل أي ارشاد عن الطريق الصحيح الذي يجب ان يتبعه وإنما يجب عليه التعلم بنفسه من خلال التجربة والخطأ. وهو مجال واسع يتم البحث فيه في عدة تخصصات كنظرية اللعبة، بحوث العمليات والخوارزميات الوراثية.

المنهج الأساسي المتبع في تدريب معظم الشبكات العصبية هو التعلم الموجّه، ويمكن الاستفادة من التعلم غير الموجه لعمل توصـيف مبدأي على مدخلات النموذج، إلا أنه قلت الدراسـات التي عالجت مشـكلة تدريب الشبكات العصبية باستخدامه (Pouyanfar et al, 2018, P.1-36).

إحدى الأهداف الرئيسية التي يجب الوصول اليها عند اتباع التعلم الموجّه هو القدرة على التعميم. والتعميم هو قدرة الشبكة العصبية على إعطاء مخرجات دقيقة عند استخدام مدخلات لم تكن قد تدربت عليها سابقاً. للوصـول إلى هذا الهدف يتم تقسـيم البيانات المتاحة إلى ثلاث مجموعات (Alpaydin, 2014, P.40) على الشكل التالي:

1) بيانات التدريب: هي البيانات الأساسية التي تدرب عليها الشبكة.
2) بيانات التحقق: تشـمل عملية تدريب الشـبكات العصـبية التجريب المسـتمر لمعلمات الشـبكة، كسـرعة التعلم، حجم الدفعة، عدد خطوات التدريب وغيرها. حيث أنه بعد كل حلقة تدريب يتم تعديل المعلمات بعد اختبارها على مجموعة بيانات التحقق. على مدى عدة حلقات تدريب تصـــبح هذه المجموعة جزء من بيانات التدريب.
3) بيانات الاختبار: الهدف منها هو اختبار فعالية الشـــبكة بالقيام بعملها على بيانات لم تتعرض لها من قبل، وذلك بعد تدريب الشبكة بشكل كامل. حيث لا يتم إعادة تعديل معلمات النموذج بعد الانتهاء من بيانات الاختبار. يحاكي استخدام هذه المجموعة من البيانات عملية اختبار النموذج على أرض الواقع.

وبهذا الأمر، تتم محاكاة عملية تعريض الشـــبكة على مجموعة بيانات جديدة لم تتم معالجتها في الســـابق وتحقيق (بشـــكل تقريبي) هدف التعميم دون الحاجة إلى اختبارها على أرض الواقع بشـــكل مباشـــر، الأمر الذي قد يكون مكلفا في بعض التطبيقات العملية.

3-1-6- خوارزميات التعلم في الشبكات العصبية الاصطناعية

تتوقف عملية التعلم على التعديل المســتمر للأوزان في الشــبكة العصــبية. حيث يجب حســاب حجم هذا التعديل لكل وزن من أوزان الشبكة بهدف تعظيم دقتها. ويتم الوصول إلى الأوزان الأكثر مناسبة على عدد مراحل. ففي البداية، يتم حساب مقياس لأخطاء الشبكة العصبية. تدعى الدالة التي تحسب الأخطاء دالة الكلفة أو دالة الخسارة وهي عبارة عن مجموع مربعات الأخطاء بين مخرجات الشبكة التقديرية والمخرجات

الفعلية. يمكننا من خلال قواعد الاشتقاق الوصول إلى المشتق الجزئي لأوزان الشبكة، والذي يعطينا دالة تربط بين الكلفة وكل وزن من أوزان الشبكة العصبية.

يكون الهدف بعد الوصول إلى هذه الدالة الحصول على اقل قيمة لها، أي أقل فرق بين مخرجات الشبكة والمخرجات الفعلية. لتحقيق ذلك تستخدم خوارزمية Gradient Descent، التي تقوم بإيجاد مجموعة المتغيرات التي تقوم بإعطاء أقل خطأ ممكن (Alpaydin, 2014, P.248).

تستخدم أيضا معها خوارزمية الانتشار العكسي Backpropagation Algorithm، والتي تقوم بحساب مصفوفة المشتقات الجزئية، حيث تعبّر هذه المصفوفة عن مقدار حساسية كل وزن من أوزان الشبكة اتجاه دالة الكلفة. وبهذا الأمر تعمل الخوارزميتين معاً على عدة دورات لتعديل هذه الأوزان والوصول إلى القيم التي تعطي اقل خطأ (Nielsen, 2015, P.53).

تمر عملية التدريب بالمراحل الأربعة التالية:

1) إعطاء قيمة عشوائية لجميع أوزان الشبكة.
2) مرحلة الانتشار الأمامي: يتم فيها إدخال البيانات ومعالجتها في جميع الطبقات (باستخدام القيم الحالية للأوزان) للوصول إلى قيمة المخرجات.
3) مرحلة الانتشار العكسي: تحسب فيها مصفوفة المشتقات الجزئية، ويعدل كل وزن بمقدار بسيط حسب المشتق الجزئي له.
4) إعادة المراحل من 1 إلى 3 بشكل كافي للوصول إلى دقة شبكة ملائمة.

3-2- المبحث الثاني - التعلم المعزّز Reinforcement Learning

تمهيد

التعلم المعزّز هو إحدى مجالات التعلم الآلي الذي يقوم على تدريب نموذج لاتخاذ مجموعة من القرارات في بيئة محددة، وذلك للوصول إلى هدف ما، كإيجاد أفضل طريق أو أفضل سلوك في وضع معين. ويتم ذلك عن طريق تحفيزه بمكافأة عند اتخاذه للقرارات الصحيحة والاقتراب من الهدف. وهو يختلف عن التعلم الموجّه حيث لا يوجد بيانات لتدريب العميل وإنما يكون عليه التجريب والتعلم بنفسه من خلال التعامل مع البيئة. خلال عدة دورات من التعلم والخطأ العشوائي يقترب العميل تدريجياً من السياسة المثلى (Sutton and Barto, 2018, P.7).

السياسة المثلى هي سلسلة القرارات التي إذا اتخذها العميل بشكل متتابع ستؤدي إلى تعظيم مكافأته، وأن الحصول على السياسة المثلى هو الهدف النهائي من عملية التدريب. ففي أنظمة التعلم الموجّه يولّد النموذج المستخدم (كالشبكات العصبية) تنبؤ أو اقتراح. فتكون مَهمة مستخدم هذا النموذج إيجاد وبناء السياسة المثلى يدويا بناءً على هذه التنبؤات. من الناحية الأخرى تقتصر مهمة مصمم مشكلة التعلم المعزّز على تحديد آلية المكافأة من دون إعطاء النموذج أي تلميح أو إيحاء عن كيفية حل المشكلة، ويكون على العميل إيجاد السياسة المثلى.

حيث يتم تصميم المكافأة لتحفز العميل بطريقة تساعده على الاقتراب من هدف النموذج ككل. على سبيل المثال عند تدريب نظام الإيقاف المؤتمت في السيارات الذكية، يمكن أن تكون المكافأة عبارة عن دالة تستخدم الفرق في المسافة بين المكان الفعلي والمكان المطلوب. وفي بعض الحالات يمكن أن تكون المكافأة ذات قيمة سالبة (أي تعامل كعقاب للعميل) وتستخدم لتدريب النموذج على الابتعاد عن قرارات معينة.

3-2-1- تطبيقات التعلم المعزّز

أستُخدِم التعلم المعزّز مجال التداول المؤسساتي، حيث تقوم مؤسسات الاستثمار أحيانا بالدخول بعمليات استحواذ أو التخلص من مجموعات كبيرة من الأصول المالية في فترة قصيرة. يمكن أن تؤثر عمليات البيع والشراء في سعر هذه الأصول في السوق خلال نفس الفترة، الامر الذي سيؤدي إلى التأثير في ربحية هذه العمليات. حيث يعرف هذا الأمر بمصاريف التنفيذ. أستُخدِم التعلم المعزّز في هذه الحالة للوصول لأفضل استراتيجية للاستحواذ أو التخلص من هذه الأصول وأدى إلى تخفيض مصاريف التنفيذ بنسبة 25% إلى 40% مقارنة مع الطرق التقليدية كتوزيع قيمة العملية بمقدار متساوي على عدة أيام (Bertsimas and Lo, 1998, P.1).

كما أستُخدِم في مجال إدارة محفظة من أصلين ماليين، يمثل كل منهما مؤشر أسهم ومؤشر سندات، حيث يكون متاح للعميل اتخاذ قرارات بخمس توزيعات مختلفة لكل من هذين الأصلين ضمن المحفظة في كل فترة تداول، وتمثل الأوضاع بحالة من حالتين لكل واحد من الأصول في كل فترة، أي تأخذ واحدة من 4 احتمالات. على الرغم من بساطة القرارات والأوضاع حقق النموذج نتائج إيجابية على بيانات فعلية (Pendharkar and Cusatis, 2018, P.1).

3-2-2- صفات التعلم المعزّز

يتصف التعلم المعزّز بالصفات الثلاثة التالية (Sutton and Barto, 2018, P.2):

1- يعتبر دارة مغلقة، أي أن القرارات التي يتخذها العميل تساهم بمدخلات المراحل اللاحقة في التعلم.

2- لا يتم اخبار العميل ما هي القرارات التي يجب أن يتخذها، بل عليه أن يجرب ويكتشف بنفسه القرارات التي تعطي المكافأة الأعظم.

3- يمكن أن يكون هنالك تأثير للقرارات على أكثر من مرحلة في المستقبل، أي أن المكافآت التي يكتسبها العميل يمكن أن تكون على أكثر من مرحلة، ومبنية على سلسلة معينة من القرارات. الأمر الذي يتطلب شيء من التخطيط والرؤية المستقبلية.

3-2-3- التطور التاريخي للتعلم المعزّز

وصل التعلم المعزّز إلى ما عليه الآن من خلال التطورات الحاصلة في مجالين، المجال الأول هو التحكم الأمثل Optimal Control والذي بدأ العمل عليه في نهاية الخمسينات، حيث يقوم على تصميم نموذج تحكم يهدف إلى تعظيم أو تخفيض مقياس معين في سلوك نظام ديناميكي على مر الوقت. يستخدم هذا المجال مفهومي وضع النظام الديناميكي ودالة القيمة لتعريف معادلة تدعى بمعادلة بيلمان Bellman Equation وتدعى مجموعة الطرق التي تقوم بحل مشاكل التحكم الأمثل بالبرمجة الديناميكية (Bellman, 1957, P.679).

أما المجال الثاني فيقوم على فكرة التعلم من خلال التجربة والخطأ والتي وضحها (Thorndike, 1970, P.244) من خلال تجاربه على الحيوان حيث وجد أنه كلما ازدادت / قلّت درجة الرضا أو عدم الارتياح اتجاه حالة ما، كلما ازدادت / قلّت شدة العلاقة اتجاهها. وسماها قانون التأثير، الذي يعد المبدأ الأساسي الذي بنيت عليه نظريات التعلم اللاحقة (Hull, 1930, P241).

بالإضافة إلى ذلك يقوم هذه المجال على فكرة التعزيز التي اكتشفها العالم الروسي بافلوف (Pavlov, 1997, P.936) وتقوم على ردود الفعل المشروطة، ووصفها على أنها التقوية الحاصلة في نمط سلوك الحيوان عندما يتعرض إلى محفز في علاقة زمنية مع محفز آخر. يشترط في تعريفه أن يكون التأثير ممتد لفترة بعد تقديم الحافز.

خلال فترة الخمسينات تم بناء العديد من الآلات الميكانيكية التي وضحت إمكانية التعلم من خلال التجربة والخطأ، فعلى سبيل المثال، وضعت إحدى الآلات لاكتشاف الطريق الأمثل الذي يؤدي إلى نهاية متاهة، واستخدمت المفاتيح الكهربائية لتذكر هذا الطريق (Walter, 1950, P.42). هذا الامر ساعد الوسط العلمي بتحويل هذه المعرفة إلى أجهزة الحاسوب الرقمية. إلا أن فترة الستينات والسبعينات قد لاقت الكثير

من التشوش والاختلاط بين التعلم المعزّز والتعلم الموجّه حيث أن الفرق بين النوعين لم يكن موضّح بشكل صريح بعد. مما أدى إلى تراجع الأبحاث في مجال التعلم المعزّز.

العمل الأساسي الذي كان مسؤولا عن إحياء مجال التعلم من التجربة والخطأ قدمه (Klopf, 1975, P.11-13) حيث لاحظ غياب العديد من الجواب المهمة للتكيف السلوكي في الأبحاث التي ركزت على التعلم الموجّه، الأمر الذي تفتقده هذه الأبحاث هو الدافع للتحكم في بيئة النموذج وتوجيهها نحو الغايات المرغوبة وبعيداً عن الغايات غير المرغوبة. كما لحق هذا العمل العديد من الأبحاث التي كرست لتوضيح أن التعلم المعزّز هو بالفعل مختلف عن التعلم الموجّه (Sutton and Barto, 1985, P.135) (Barto and Anandan, 1985, P.360).

فيما بعد تم تطوير خوارزمية Q-Learning (Watkins, 1989) لحل مشكلة التعلم المعزّز، والتي دمجت بين مجال التحكم الأمثل والتعلم من خلال التجربة والخطأ. ومنذ ذلك الوقت نمت أبحاث التعلم المعزّز بشكل هائل وتوسعت إلى مجالات عديدة كاستخدامها مع الشبكات العصبية، حيث أدى نجاح برنامج TD-Gammon (Tesauro, 1994, P.215) المخصص للعب لعبة الطاولة إلى رفع مستوى الاهتمام في هذا المجال بشكل هائل.

3-2-4- عمليات اتخاذ القرار الماركوفية Markov Decision Processes

تدعى أيضا بعمليات ماركوف وهي عبارة عن إطار تنظيمي لعملية اتخاذ القرارات التسلسلية. في ضمن هذا الإطار إن عملية اتخاذ القرار لا تؤثر بالمكافآت الآنية (التي سيتم الحصول عليها في وقت القرار) فحسب، بل ستؤدي إلى تغيير الحالات اللاحقة أيضاً، وبالتالي ستتغير المكافآت المستقبلية حسب هذه الحالات. أي أن عمليات ماركوف تنطوي على المكافآت المؤجلة والحاجة إلى المقايضة بينها وبين المكافآت الآنية. ويحتوي هذا الإطار على الخاصية الماركوفية التي تعني أن الحالات السابقة لا تؤثر بالحالات المستقبلية.

يتمثل دور عمليات اتخاذ القرار الماركوفية بتعريف البيئة، أما التعلم المعزّز فيقدم الطرق للتعامل مع هذه البيئة والتعلم منها للوصـــول إلى الســياســـة الأمثل. يمكن ان تكون البيئة معرفة بالكامل لكن يمكن تعميم المشـكلة لتتضــمن تعريف جزئي للبيئة (Spaan, 2012, P387). وهي تتألف إضــافة إلى العميل من المكونات التالية (Bäuerle and Rieder, 2011, P.1):

1) أوضاع أو حالات البيئة (State (S)، وهي المعلومات المتاحة للعميل في المرحلة الزمنية t، وتضمن جميع البيانات الضرورية التي يحتاجها العميل لاتخاذ القرار في هذه المرحلة.
2) مجموعة القرارات (Action (A)، وهي القرارات المتاحة للعميل في كل حالة.
3) دالة احتمالات الانتقال $PR(s_t + 1 \mid s_t, a_t)$ تعطي احتمال الوصـــول إلى الحالة التالية St+1 عندما يكون العميل في الحالة s ويتخذ القرار a، وتعبّر عن آلية عمل البيئة.
4) دالة المكافأة $R(s, a)$ او $R(s)$، وهي العائد المتوقع الذي سيحصل عليه العميل إذا كان في الحالة s واتخذ قرار a، او قيمة (مكافأة) الحالة s. وهي عبارة عن جزئين، الأول هو المكافأة الحالية، والثاني هو المكافآت المستقبلية المجمعة المخصومة.

يوجد نوعين من عمليات ماركوف، الأول تكون فيه الحالات والقرارات محدودة والأخر تكون الحالات أو القرارات مستمرة وغير محدودة. والنوع الثاني هو الأكثر مناسبة لنمذجة المشاكل المعقدة كمشكلة تداول الأصول المالية على سبيل المثال التي يمكن أن تمثل الحالات فيها بأسعار هذه الأصول، والتي تأخذ قيم حقيقية لا نهائية. هذا التعقيد سيتسبب بمشكلة عند تدريب العميل إلا أنه يمكن تفاديها بالاستعانة بمجالات التعلم الآلي الأخرى كالشبكات العصبية كما هو الحال في التعلم المعزّز العميق.

3-2-5- خوارزميات التعلم المعزّز

في التعلم المعزّز يوجد العديد من الخوارزميات لحل عمليات ماركوف، تصنف هذه الخوارزميات حسب وجود نموذج للبيئة إلى (Sutton and Barto, 2018, P.12):

1) خوارزميات قائمة على نموذج Model-based algorithms: تستخدم عندما تكون العمليات الداخلية في البيئة معروفة، أي أن العميل لديه معرفة دقيقة عن الحالة الذي سيصل اليها والمكافأة التي سيحصل عليها إذا اتخذ أي قرار. أي دالة احتمالات الانتقال تعطي الاحتمالات الحقيقية للانتقال من الحالة s عند اتخاذ القرار a وليس القيم التقديرية المحسوبة.
2) خوارزميات دون نموذج Model-free algorithms: وهو عكس الحالة السابقة، عندما تكون البيئة معقدة وعملياتها الداخلية غير معروفة.

كما يمكن أن تصنف خوارزميات التعلم المعزّز حسب هدف الخوارزمية إلى (المرجع السابق نفسه، ص. 342):

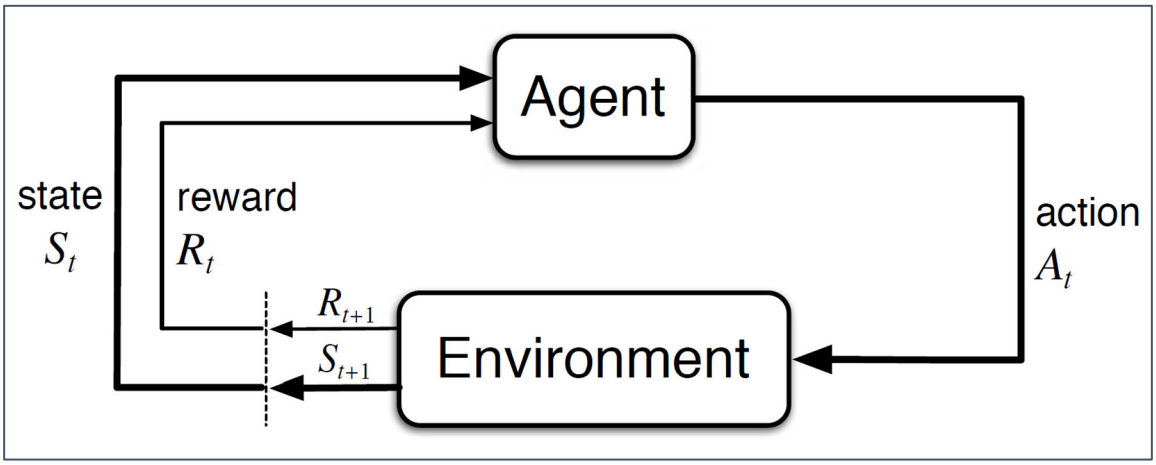
1) خوارزميات التنبؤ Prediction Algorithms: الهدف منها هو الوصول إلى أحد معالم البيئة كدالة المكافأة والتي تعد أهمها، وذلك من خلال سياسة معينة متوفرة (ليس بالضرورة أن تكون الأمثل) كمدخلات للخوارزمية.
2) خوارزميات التحكم Control Algorithms: هدف خوارزميات التحكم هو استكشاف البيئة واستنتاج السياسة الأمثل منها.

معظم المشاكل على أرض الواقع تتطلب خوارزميات من تحكم من دون نموذج Model-Free Control Algorithms، أهم هذه الخوارزميات:

1) تحكم مونتي كارلو Monte Carlo Control
2) الوضع-القرار-المكافئة-الوضع- القرار State Action Reward State Action (SARSA)
3) تعلم Q (Q Learning)

3-2-6- الية تدريب عميل التعلم المعزّز

الشكل 3 - التفاعل بين العميل والبيئة



Source: (Sutton and Barto, 2018, P.54)

في جميع خوارزميات التحكم التي تعمل من دون نموذج، يقوم العميل بالتفاعل مع البيئة من خلال عدة خطوات زمنية متتابعة بهدف الوصول إلى السياسة الأمثل، ليستطيع فعل ذلك يمر بأربع مراحل كما يلي (Sutton and Barto, 2018):

1) إعطاء تقديرات مبدئية:

لتحقيق الهدف الأساسي من التعلم المعزّز يجب أولا الوصول إلى تقدير عالي لدالة المكافأة. يمكن أن تمثل دالة المكافأة كجدول كل سـطر فيه يعبر عن الحالة وكل عامود هو القرار، وقيم الجدول هي أرقام حقيقية تعبر عن درجة أهمية زوج الحالة والقرار معاً. في البداية تعطى قيمة صـفر لجميع قيم الجدول.

2) اختيار القرار:

يختار العميل قراراً من الحالة الأولية التي يبدأ فيها، والتي سـتنقله إلى الحالة الجديدة، والتي منها سيقوم باتخاذ قرار جديد، وهكذا.

على العميل اختيار القرار بحيث يكون قادرا على التعرف على قيم جميع القرارات وفي نفس الوقت تعظيم مجموع مكافآته. تأتي هنا أهمية المقايضة بين وضعي الاستكشاف والاستغلال Exploration and Exploitation:

أ. في وضع الاستكشاف يقوم العميل بالتعرف على أكبر قدر ممكن من قيم مكافآت ازواج الحالة والقرار، وذلك عن طريق اختياره للقرار في كل حالة بشكل عشوائي.

ب. وفي وضع الاستغلال، يسعى العميل إلى اتخاذ القرارات التي تؤدي إلى حصوله على أكبر مكافأة، لكن ليس بالضرورة أن تكون هذه المكافأة هي الأمثل، لأنه من الممكن أن يكون هنالك قرارات لم يكتشفها بعد قد تؤدي إلى مكافأة أعلى.

تتم عملية المقايضة من خلال استراتيجية $\varepsilon - Greedy$ والتي تقوم على استخدام وضع الاستكشاف في الخطوات الزمنية الأولى للتدريب ومن ثم الانتقال تدريجياً إلى وضع الاستغلال في الخطوات اللاحقة.

3) الحصول على تغذية راجعة من البيئة:

بعد أن قام العميل باتخاذ القرار، سينتقل إلى الحالة الجديدة وسيحصل على المكافأة بالإضافة إلى معلومات الحالة التي أدى هذا القرار إليها.

4) تحسين التقديرات:

في هذه المرحلة يقوم العميل بتعديل جدول المكافأة بقيمة تدعى العائد، يتألف العائد من ثلاث مكونات:

أ. القيمة المقدرة السابقة الموجودة في جدول المكافأة والتي أخذت قيمة صفر في المرحلة الأولى.

ب. المكافأة التي حصل عليها في المرحلة السابقة.

ج. القيمة المتوقعة للمكافآت المستقبلية (من جدول المكافأة) مخصومة بمعامل خصم.

تعاد المراحل من 2 إلى 4 على عدة حلقات حتى يصل العميل إلى أعلى درجات من التقدير لجدول المكافأة، ومنه إلى السياسة الأمثل.

3-2-7- التعلم المعزّز العميق Deep Reinforcement Learning

التعلم المعزّز العميق هو مجال التعلم الآلي الذي يستخدم الشبكات العصبية الاصطناعية في إطار التعلم المعزّز. ويصنف عادة ضمن التعلم المعزّز بسبب تشابه الهدف بينهما وهو تعلم أفضل طريقة لتحقيق هدف معين. تتميز الشبكات العصبية الاصطناعية المستخدمة بأنها عميقة، أي تحوي على عدة طبقات مخفية متسلسلة يمكن ان تتجاوز 100 طبقة (Szegedy et al., 2017)، يتم في كل واحدة من هذه الطبقات التحويل غير الخطي لمخرجات الطبقة السابقة، وتؤدي سلسلة التحويلات إلى تعلم مستويات مختلفة من التجريد في البيانات (Olah, 2017).

في التعلم المعزّز عندما تكون البيئة معقدة وعدد حالاتها كبير أو غير نهائي (سعر سهم، أو المسافة بين السيارة التي تتعلم القيادة الآلية والاشياء في محيطها مثلاً)، يكون من الصعب على العميل استكشاف جميع الحالات الممكنة، وبالتالي تقديره لدالة (جدول) المكافأة سيقتصر على الحالات التي واجهها خلال تدريبه فقط، أي انه إذا واجهته حالة مختلفة بشكل بسيط عن حالة قد تدرب عليها (على سبيل المثال كأن يكون سعر السهم 130.2 بدل 130.1، أو أن تكون المسافة بين سيارة العميل وبين عامود الكهرباء 5.5 متر بدل من 5.7)، لن يعطي النتائج المرجوة. هنا يأتي دور التعلم العميق، حيث تعتبر الشبكات العصبية مقربات دالة، أي أنه إذا تم إنشاؤها وتدريبها بشكل صحيح ستكون قادرة على حل محل دالة أخرى وستعطي نتائج قريبة للدالة الفعلية (Nielsen, 2015, P.128). عوضاً عن تدريب العميل للتعرف على جميع حالات البيئة التي يمكن أن تكون لا نهائية، يمكننا تدريب شبكة عصبية باستخدام بيانات فعلية لتحل محل دالة المكافأة، حيث تأخذ هذه الشبكة وصف الحالة كمدخلات، وتعطي قيمة هذه الحالة أو المكافأة المقدرة. ونفس الأمر ينطبق في حال تقدير قيمة زوج الحالة والقرار معاً. وبعد ذلك تستخدم المكافآت لاستخراج السياسة المثلى.

أحد المجالات التطبيقية للتعلم المعزّز العميق هو التداول في الأســـواق المالية حيث يمكن تدريب عميل التعلم المعزّز على الاختيار بين إبقاء الاســتثمار، بيعه أو شــراءه، حيث تحســب قيمة المكافأة في هذه الحالة حســب العائد أو الخســارة المحققة من التداول. وبالتالي يشــجع العميل على البحث والتعرف على الأوضـــاع في البيئة التي يكون عليه اتخاذ قرار فيها، وما هو نوع القرار الذي يعظم عائده. تقوم شــركة Equbot بتطوير هذه النماذج واســتخدامها في مؤســســات اســتثمار فعلية (Beverungen, 2019, P.77).

3-2-8- معالجة مشكلة إدارة المحفظة المستمرة باستخدام التعلم المعزّز العميق

النموذج المستخدم في هذا البحث مبني على نموذج (Jiang et al., 2017, P.15) الذي وظف التعلم المعزّز إضـــافة إلى التعلم العميق، حيث قام باختبار أداء ثلاث أنواع من الشــبكات العصــبية CNN، RNN، LSTM. حقق كل نوع من هذه الشـبكات نتائج مختلفة، وكان النموذج الذي يحوي شـبكة CNN هو الأفضـل وسطياً من حيث معايير نسبة شارب، قيمة المحفظة النهائية ونسبة الخسارة العظمى، لذلك سيتم اعتماده في هذه الدراسة. حيث يتفاعل كل من التعلم المعزّز والشبكة العصبونية ضمن النموذج على الشكل التالي:

1) التعلم المعزّز ضمن نموذج الدراسة

في إطار التعلم المعزّز يمثّل العميل دور مدير المحفظة بقيامه باختيار أوزان أصـــول المحفظة في بيئة السوق المالي. تتضمن هذه البيئة لكل من الأصول المالية المتاحة للتداول:

1. جميع المعلومات الأساسية المتوفرة عن هذه الأصـول كمبيعات الشـركة، نسـبة العائد على رأس المال، والوضع الاقتصادي للقطاع التي تعمل به الشركة.
2. مجموعة المؤشرات الفنية كخطوط الدعم والمقاومة، مؤشر عزم القوة النسبية، والمتوسط المتحرك، جميعها مستخلصة من الأسعار الحالية والتاريخية للأصل المالي.

وبالتالي هنالك كميات هائلة من المعلومات التي إذا اســتخدمت جميعها لبناء بيئة مماثلة للواقع، ســيكون من الصــعب تدريب العميل باسـتخدامها. لذلك سـيتم إتّباع فلسـفة التداول الفني (Kirkpatrick et al., 2006) (Lo et al., 2000, P.1705) والتي تقوم على مبدأ أن جميع المعلومات المتاحة عن الأصـل تنعكس في سعر الأصل، أي أن استخدام السعر التاريخي للأصل سيكون كافيا لتمثيل أوضاع البيئة.

أما قرارات العميل فهي عبارة عن أوزان أصـول المحفظة التي يجب أن تكون في الفترة الزمنية التالية، لكن هذه الأوزان مرتبطة بشـكل غير مباشـر مع أوزان أصـول المحفظة في الفترة السـابقة، أي أنه يجب على العميل اخذها بعين الاعتبار عند اتخاذه للقرار، وبالتالي يجب إدخال أوزان الفترة الســـابقة مع أوضـــاع البيئة.

حتى يســـتطيع العميل تحقيق هدفه بتعظيم قيمة المحفظة، يجب اســـتخدام دالة مكافأة لقياس أداء العميل وفعالية القرارات التي يتخذها من أجل تخفيض احتمال اختياره للقرارات التي لا تســـاهم بتحقيق هدفه ورفع احتمال اتخاذ القرارات التي تزيد قيمة المحفظة النهائية. يمكننا قياس المكافأة الجزئية لكل قرار من خلال احتساب عائد المحفظة المحقق عند اتخاذ القرار. بما أنه يتم تدريب العميل على عدد من الحلقات وتتكون كل حلقة من عدة فترات تداول، فيجب أخذ وســـطي المكافآت الجزئية للفترات (عن طريق القانون التالي) من أجل الوصـــول إلى العائد الكلي لحلقة التدريب. وفي نهاية كل حلقة تدريب يتم تعديل معلمات النموذج ليصبح العميل أقرب من السياسة المثلى.

$$R = \frac{1}{T}\sum_{t=1}^{T} r_t$$

$\boldsymbol{R}$: العائد الكلي لحلقة التدريب.

$\boldsymbol{T}$: اجمالي فترات التداول ضمن حلقة التدريب.

$\boldsymbol{r_t}$: عائد فترة التداول t.

وبالتالي يعرف إطار التعلم المعزّز من خلال:

1) أوضاع البيئة s_t: والمؤلفة من شقين الأول هو سعر الأصل المالي في الفترة الحالية $\boldsymbol{v_t}$ والشق الثاني هو أوزان المحفظة في الفترة السابقة w_{t-1}.
2) القرارات a_t: وهي الأوزان التي اختارها العميل لأصول المحفظة w_t.
3) المكافأة R: المعرفة بالمعادلة السابقة وهي المؤشر الذي يرشد العميل نحو السياسة المثلى.

2) الشبكة العصبية الاصطناعية ضمن نموذج الدراسة

تؤدي الشبكة العصبية دور عميل التعلم المعزّز. حيث تأخذ وضع البيئة كمدخلات لها، والتي تتكون من أسعار الاغلاق، الأسعار العليا والأسعار الصغرى لجميع الأصول المتداولة، إضافة إلى أوزان الفترة السابقة التي تضاف لاحقاً في إحدى الطبقات المخفية. تعطي الشبكة (قرار العميل) متجه بأوزان الأصول التي يجب أن تكون عليه في الفترة التالية (Jiang et al., 2017, P.14).

تتكون الشبكة العصبية المستخدمة من الطبقات التالية:

1) طبقة مدخلات
2) أربع طبقات مخفية:
   a. ثلاث طبقات تلافيفية
   b. طبقة من نوع softmax
3) طبقة مخرجات

الطبقة التلافيفية convolutional layer هي نوع من الطبقات المناسبة للاستخدام مع البيانات التسلسلية والصور. تقوم بتطبيق عملية تلافف على مدخلاتها لاستخراج ميزات معينة في البيانات ومن ثم تقوم بإعطاء المخرجات للطبقة التي تليها، والتي بدورها تتعلم أنماط أبسط من الميزات التي تلقتها من الطبقة السابقة، وهكذا. تؤدي هذه العملية إلى قيام الشبكة بتعلم الملامح الأساسية في البيانات من خلال تجزيئها إلى أقسام سهلة التعرف على عدة طبقات.

في تصنيف الصور على سبيل المثال تقوم الطبقات الأولى بالتعرف على الحواف، البنية والانماط ضمن الصورة. والطبقات التي تليها تتعلم أجزاء الأشياء في الصورة ومن ثم كامل الأشياء ضمنها (Olah et al., 2017, P.1). على الرغم من ان بيانات أسعار الأصول ليست صور، إلا أن تنظيمها على شكل ثنائي البعد كالصور –من خلال استخدام جدول تمثل أسطره نوع الاصل مالي، أعمدته الفترات الزمنية، وقيمه أسعار الأصول– يتيح قابلية استخدامها مع الشبكات التلافيفية.

تستخدم طبقة softmax للحصول على توزيع احتمالي، حيث تتكون من عامود واحد من العصبونات المتصلة بشكل كامل مع الطبقة التلافيفية التي تسبقها، تقوم بتلقي مخرجات الطبقة التلافيفية وتحويلها بحيث يصبح مجموع جميع قيم العصبونات فيها واحد، وبالتالي تعبر مخرجاتها عن وزن كل أصل ضمن المحفظة مع إبقاء شرط الأوزان المعرف في قسم التمثيل الرياضي لمشكلة إدارة المحافظ المستمرة محققاً (Nielsen, 2015, P.70).

يدرَّب عميل التعلم المعزّز العميق باستخدام طريقة Deterministic Policy Gradient، حيث تعرّف السياسة التي يسعى إلى تعلمها من خلال معلمات الشبكة العصبية التي تتضمن أوزان عصبونات الشبكة، ومقياس الأداء المستخدم خلال عملية التعلم هو دالة المكافأة. تبدأ عملية التعلم بإعطاء قيم عشوائية لمعلمات الشبكة، ومن ثم حساب المشتقات الجزئية لها وتحديثها بشكل مستمر باستخدام نسبة مئوية تدعى سرعة التعلم وهي احدى معلمات النموذج الكلي.

لرفع فعالية وسرعة التدريب يتم القيام بعملية التحديث للمعلمات بعد تدريبها على دفعات صغيرة –Mini batches من البيانات، عوضاً عن كامل الأسعار التاريخية.

3-2-9- معلمات النموذج الكلي Hyperparameters

هي معلمات مختلفة عن معلمات الشـــبكة العصـــبية، تتحكم معلمات النموذج الكلي بفعالية تدريب وأداء النموذج ككل. تلعب عملية التجربة دورا أساسيا في تحديد هذه المعالم، حيث بعد أن يتم تحديدها يقيم أداء النموذج ومن ثم تعدل المعلمات بشكل بسيط ويعاد التقييم، وهكذا حتى نصل إلى معلمات النموذج الكلي ذات الفعالية الأعلى. من اهم معلمات النموذج الكلي:

1) سرعة التدريب:

تحدد حجم التعديل الحاصل في معلمات الشبكة العصبية. إن اختيار سرعة تدريب مرتفعة قد يؤدي إلى تخفيض الوقت اللازم لتدريب النموذج، إلا أنه من الممكن أن يجعل العميل غير قادراً على الوصول إلى الحد الأدنى لدالة المكافأة، وبالتالي ستكون النتائج غير مثالية. لذا يجب اختيار سرعة تدريب متوسطة لأخذ هذين الأمرين بعين الاعتبار.

2) حجم النافذة:

مدة فترة التداول (عدد الأيام) المســـتخدمة في عينة التدريب الواحدة، كلما كان حجم النافذة أكبر كلما ازدادت البيانات المتوفرة للعميل حتى يقوم باتخاذ قراره. ليس من الجيد أن يكون حجم النافذة كبيراً لأن هذا الامر ســـيؤدي إلى دخول كثير من المعلومات القديمة ذات الأهمية والتأثير المنخفض على حالة الأصل في المستقبل. تتصف النوافذ بكونها متحركة ومتداخلة مع بعضها.

3) حجم الدفعة:

عدد النوافذ في حلقة التدريب الواحدة، يمكن أن تكون النوافذ مختارة بشكل عشوائي من مجتمع العينة، أي أن تكون غير متتالية، كما يمكن أن تكون متتالية وغير عشـــوائية. بعد القيام بتدريب كامل الدفعة يتم حساب متوسط أداء عيناتها ومن ثم تعدل معالم الشبكة العصبية.

4) خطوات التدريب:

هي عدد مرات تكرار عملية التدريب. يجب ألا يكون هذا الرقم قليلاً بحيث لا يصـــل العميل إلى الحد الأدنى من دالة المكافأة، وألا يكون كبيراً كي لا يعاني النموذج من مشـــكلة overfitting والتي تعني عدم قدرته على التعميم.

كما يوجد معلمات أخرى كعدد الأصـول الداخلة المسـتخدمة، نسـبة بيانات الاختبار من اجمالي البيانات، وغيرها.

الفصل الرابع – الدراسة التطبيقية

مقدمة

في المبحث الأول من هذا الفصل سيتم التعرّف على الأسواق المالية التي سيتم تطبيق الدراسة عليها، وهي سوق دمشق للأوراق المالية وبورصة عمّان. وفي المبحث الثاني سيتم التعرف على بيانات الدراسة، من طريقة اختيار الأسهم وجمع ومعالجة البيانات واختيار الفترات وطرق تقييم النماذج التي تم استخدامها والنتائج التي تم التوصل اليها في كل سوق مالي. في المبحث الثالث والأخير سيتم إجراء اختبار إحصائي لفرضيات البحث مناقشة النتائج، والاختتام بالتوصيات والافاق المستقبلية للبحث.

4-1- المبحث الأول – لمحة عن الأسواق المالية في الدراسة

تمهيد

تعد الأسواق المالية بشكلها الحالي حديثة العهد بالمقارنة مع الأسواق الأخرى حيث أن ضخامة الاستثمارات المالية فيها رفع من درجة أهميتها وأدى إلى تطورها بشكل هائل، سواء من حيث التنظيم أو من حيث الإمكانيات والتسهيلات المتاحة للمتعاملين فيها.

وقد تزايد الدور الاقتصادي للأسواق المالية وأصبحت تحتل مركزا حيويا في النظم المعاصرة التي تقوم على أساس التعامل الحر، حيث يتم من خلالها تجميع المدخرات وتوجيهها نحو الاستثمار الذي يؤدي إلى زيادة السلع والخدمات. فهي تمثل حلقة من حلقات تطور النظام المصرفي وانعكاساً لنظمه الاقتصادية والمالية، فالسوق المالي هو بمثابة وسيط بين الأفراد والبنوك والمؤسسات الأخرى.

4-1-1- لمحة عن سوق دمشق للأوراق المالية

بهدف دفع عجلة الاقتصاد السوري وسعياً لتحسين البنية المالية السورية، أصدر المرسوم التشريعي رقم 55 لعام 2006 الذي أحدث بموجبه أول سوق لتداول الاوراق المالية في الجمهورية العربية السورية تدعى "سوق دمشق للأوراق المالية". تتمتع السوق بشخصية اعتبارية مستقلة واستقلال مالي وإداري وترتبط بهيئة الأوراق والأسواق المالية السورية وتعمل تحت اشرافها، ويحدَّد موقع السوق في مدينة دمشق.

تسعى السوق إلى توسيع النشاط الاقتصادي من خلال توفير البيئة الملائمة لتسهيل عملية استثمار الأموال وتوظيفها بالإضافة إلى تأمين التمويل اللازم للشركات. يتم ذلك عن طريق ترسيخ أسس التداول السليم والواضح للأوراق المالية ووضع إجراءات وأساليب عمل كفيلة بضمان تداول كفؤ وشفاف، والتأكد من عدالة متطلبات الإدراج وتوفير قواعد وإجراءات سليمة وسريعة لمركز المقاصة والحفظ المركزي (كنعان، 2009، ص.63).

بدأ التداول في السوق بتاريخ 10 اذار 2009 متيحاً بذلك وسيلة جديدة للاستثمار وتعبئة المدخرات للأفراد من جهة، وامكانية النمو والتوسع لمنشئات الأعمال والشركات من جهة اخرى، وذلك من خلال توفير رأس المال اللازم في حال اتخاذ هذه المنشئات قرار التحويل إلى شركات مساهمة. كما تسهل السوق عملية جذب الاستثمارات الخارجية، الأمر الذي يؤدي إلى دفع الاقتصاد السوري إلى الأمام وتحسين الوضع المعيشي لجميع أفراد المجتمع.

4-1-2- توزع الشركات المدرجة في دمشق للأوراق المالية

تشكل البنوك الخاصة الجزء الأكبر من الشركات المساهمة المدرجة في سوق دمشق للأوراق المالية. حيث وصل عددها في الوقت الراهن إلى 14 بنك، ومن ثم يأتي بعدها شركات التأمين (6 شركات) يليها شركتين صناعيتين، شركتي خدمات، شركتي اتصالات، وشركة زراعية واحدة، ليصل العدد الاجمالي إلى 27 شركة مدرجة.

4-1-3- لمحة عن بورصة عمّان

في بداية السبعينات، بدأت فكرة إيجاد سوق أردنية منظمة للأوراق المالية بالتبلور، حيث أن الشركات المساهمة الاردنية كانت موجودة منذ الثلاثينات (بداية بالبنك العربي عام 1930 ومن ثم شركة التبغ والسجائر الأردنية عام 1931 وشركة الكهرباء الأردنية عام 1938). لكن في ذلك الوقت كانت عمليات بيع وشراء الأوراق المالية تتم لدى تجار ومؤسسات عقارية لا تحكمها قوانين أو انظمة محددة.[3] صدر القانون المؤقت رقم /31/ في عام 1976 الذي يقضي بإنشاء ما كان يدعى بسوق عمّان المالي كمؤسسة مالية تهدف إلى تنظيم ومراقبة اصدار الأوراق المالية والتعامل بها. حيث شكلت أول لجنة إدارة لهذا السوق بتاريخ 1977/03/13 وباشر السوق أعماله في 1978/01/01.

[3] https://www.ase.com.jo/ar/nbdht-n-albwrst/hwkmt-shrkt-bwrst-man/swq-ras-almal, retrieved on July 2, 2021

تبنَّت الحكومة الأردنية في عام 1997 سياسة إصلاح شاملة لسوق رأس المال تقوم على أساس تطوير ما تم إنجازه في السنوات العشرين السابقة وذلك سعياً للارتقاء إلى المعايير الدولية في مجال تنظيم الأسواق المالية. وتتمثل أهم ملامح التوجه الجديد في إحداث تغييرات مؤسسية في سوق رأس المال واستخدام أنظمة التداول والتسوية والتقاص الإلكترونية وإزالة كافة معوقات الاستثمار وتقوية الرقابة على سوق رأس المال وصولاً إلى أعلى مستوى من الشفافية وتحقيق سلامة التعامل بالأوراق المالية. وبما يتناسب مع التوجه نحو العولمة والانفتاح على العالم الخارجي.

وفي 20 شباط 2017 تم تسجيل السوق كشركة مساهمة عامة مملوكة بالكامل من قبل الحكومة الأردنية، وأصبحت تدعى "شركة بورصة عمّان" وتدار الشركة من قبل مجلس إدارة مكون من سبعة اعضاء يعينهم مجلس الوزراء ومدير تنفيذي متفرغ يتولى إدارة ومتابعة الأعمال اليومية للبورصة.

4-1-4- أهداف بورصة عمّان

تسعى بورصة عمّان إلى تحقيق الأهداف التالية (بوكرومة، 2019، ص.254):

1- ممارسة جميع أعمال أسواق الأوراق المالية والسلع والمشتقات وتشغيلها وإدارتها وتطويرها داخل المملكة وخارجها.

2- توفير المناخ المناسب لضمان تفاعل قوى العرض والطلب على الأوراق المالية المتداولة وفق أُسس التداول السليم والواضح والعادل.

3- نشر ثقافة الاستثمار في الأسواق المالية وتنمية المعرفة المتعلقة بالأسواق المالية والخدمات التي تقدمها البورصة.

4-2- المبحث الثاني - مقارنة أداء الخوارزميات في كل من سوقي دمشق وعمّان

تمهيد

يتضمن هذا المبحث توضيح للبيانات المستخدمة من طريقة جمعها، كيفية اختيار الأسهم، فترات الدراسة، توضيح المؤشرات المتبعة لتقييم أداء المحافظ، ومقارنة نتائج المحافظ.

4-2-1- جمع البيانات واختيار أسهم الدراسة

تم جمع بيانات أسعار يومية لأسهم سوقي دمشق وعمّان من الموقع الرسمي لكل سوق. حيث تتضمن أسعار الاغلاق، الأعلى، الأدنى، وحجم التداول لجميع الأسهم المتداولة من نشأتها وحتى تاريخ 2020/07/30. تحوي بيانات الأسهم على فراغات أو نقص بسبب عدم وجود تداول في بعض الفترات. تم معالجة هذا الأمر من خلال إجراء عملية Forward Filling، أي طريق ملئ النقص في بيانات اليوم من خلال نسخ بيانات آخر يوم يحوي بيانات. وهو أمر ضروري للحصول على بيانات مستمرة تعطي نتائج واقعية من النماذج المستخدمة. تم الاستعانة بلغة البرمجة Python ومكتبة Pandas للقيام بعملية الجمع والمعالجة الأولية.

نظرا لأهمية أن تكون الأسهم ذات سيولة عالية، وأهمية أن يكون هنالك عدد كافي من بيانات التداول، وبسبب صعوبة تدريب النماذج ذات عدد الاسهم الكبير، تم اختيار عدد من الأسهم (والتخلي عن باقي أسهم السوق). حيث سيستطيع عميل التعلم المعزّز توزيع الثروة المتاحة على مجموعة الأسهم المختارة في كل فترة تداول. تم إتباع المعايير التالية لاختيار مجموعة الأسهم:

1) مدى البيانات Data Range:

هو عدد الأيام بين تاريخ بداية تداول السهم وتاريخ 2020/07/30. حيث كلما ازداد عدد الأيام كلما كان من الأسهل تدريب العميل.

2) معامل التغيير لسعر السهم Price Coefficient of Variation:

يعبّر هذا المعيار عن درجة تذبذب سعر السهم، فاذا كانت هذه القيمة مرتفعة لسهم ما، سيتم تفضيله عن غيره لأنها تعني أنه يوجد العديد من عمليات التداول التي أدت إلى تغيير سعر السهم على مدى عمره. ويحسب من خلال المعادلة التالية:

$$\frac{1}{\overline{v_i}}\sqrt{\frac{\sum_{t=1}^{T}(v_{i,t}-\overline{v_i})^2}{T}}$$

حيث $v_{i,t}$ هي سعر السهم i في الفترة t، و $\overline{v_i}$ هو متوسط سعر السهم i.

3) نسبة التداول Trading Ratio:

تحسب من خلال قسمة عدد الأيام التي حدث فيها عملية تداول واحدة على الأقل على إجمالي عدد الأيام (مدى البيانات).

4) نسبة حجم التداول Volume Ratio:

وهي إجمالي حجم التداول للسهم مقسومة على مجموع حجوم التداول لجميع أسهم السوق.

5) التجربة:

من خلال إدخال عدة تركيبات من الأسهم في نموذج التعلم العميق المعزّز ومقارنة عوائد المحافظ للوصول إلى التركيبة التي تحقق اعلى عائد.

لتخفيض تحيّز الرابح[4] في النماذج المستخدمة إلى أقل حد ممكن، تم حساب جميع المعايير السابقة (عدا مدى البيانات) على جزء من البيانات هو بيانات التدريب (عوضاً عن كامل البيانات). سيتم توضيح الفترات التي وزعت عليها بيانات الأسهم في الفقرة التالية.

فيما يلي الأسهم التي تم اختيارها حسب المعايير السابقة من سوق دمشق للأوراق المالية:

الجدول 1 - اختيار اسهم سوق دمشق للأوراق المالية

Stock	Company Name	Start of Trading	End of Trading	Data Range	Price Coefficient of Variation	Trading Ratio	Volume Ratio
SGB	بنك سورية والخليج	2010/08/01	2020/07/30	3651	80%	47%	25%
SIIB	بنك سورية الدولي الإسلامي	2009/06/04	2020/07/30	4074	88%	95%	24%
QNBS	بنك قطر الوطني - سورية	2010/04/08	2020/07/30	3766	81%	91%	12%
FSBS	فرنسبنك - سورية	2011/01/09	2020/07/30	3490	63%	50%	6%
ATI	شركة العقيلة للتأمين التكافلي	2010/08/23	2020/07/30	3629	71%	61%	3%
IBTF	المصرف الدولي للتجارة والتمويل	2009/04/02	2020/07/30	4137	93%	59%	3%

حيث كان أعلى مدى للبيانات 4137 يوماً (متضمنا بيانات الاختبار) وذلك للمصرف الدولي للتجارة والتمويل. لكن بسبب وجود مصرف فرنسبنك – سورية ذات مدى البيانات الأقل (3490 يوماً) ضمن هذه المجموعة سيتم تحديد تاريخ بداية البيانات لكامل السوق بتاريخ بداية التداول لمصرف فرنسبنك – سورية (2011/01/09) ليصبح مدى البيانات مطابقاً لجميع الأسهم حتى نستطيع استخدامهم ضمن النماذج.

من حيث كل من معامل التغير لسعر السهم، نسبة التداول، وحجم التداول فقد كان ترتيب فرنسبنك – سورية الأقل من ضمن الشركات الأخرى التي تم اختيارها. على الرغم من وجود شركات ذات نسب أفضل

[4] يعني تحيّز الرابح أن نجاح النماذج قد يتأثر باختيارنا للأسهم الناجحة بشكل مسبق من خلال المعايير الخمس السابقة ولن يعطي نتائج موضوعية معزولة عن العوامل الأخرى وخاصة بأداء هذه النماذج.

الا انه تم اســتبعادها بســبب انخفاض حجم بياناتها الامر الذي ســيؤدي إلى تخفيض حجم بيانات جميع الأســهم الأخرى وبالتالي ســينخفض حجم العينة. وذلك إضــافة إلى أن بعض هذه الأســهم المســتبعدة قد أعطت نتائج غير مرضية عند التجربة على بيانات التدريب.

بالنسبة لمؤشرات أسهم بورصة عمّان فقد كانت على الشكل التالي:

الجدول 2 - اختيار اسهم بورصة عمّان للأوراق المالية

Stock	Company Name	Start of Trading	End of Trading	Data Range	Price Coefficient of Variation	Trading Ratio	Volume Ratio
PHNX	فينيكس العربية القابضة	2006/06/05	2020/07/30	5169	103%	66%	7.03%
IDMC	مجمع الضليل الصناعي العقاري	2004/01/22	2020/07/30	6034	54%	59%	5.69%
JNTH	التجمعات لخدمات التغذية والاسكان	2004/01/29	2020/07/30	6027	64%	67%	4.28%
SPIC	التجمعات الاستثمارية المتخصصة	2001/07/16	2020/07/30	6954	101%	63%	4.13%
UINV	الاتحاد للاستثمارات المالية	2001/07/16	2020/07/30	6954	103%	61%	3.41%
ULDC	الاتحاد لتطوير الاراضي	2001/07/16	2020/07/30	6954	51%	58%	3.12%
REDV	تطوير العقارات	2001/09/11	2020/07/30	6897	102%	57%	2.84%
SECO	الجنوب للإلكترونيات	2006/01/23	2020/07/30	5302	220%	57%	2.59%
JOIR	الموارد الصناعية الأردنية	2001/07/15	2020/07/30	6955	62%	63%	2.02%
TAJM	التجمعات للمشاريع السياحية	2001/08/26	2020/07/30	6913	70%	55%	1.86%
AEIV	الشرق العربي للاستثمارات المالية والاقتصادية	2001/09/06	2020/07/30	6902	159%	57%	1.72%
JOST	حديد الأردن	2001/07/15	2020/07/30	6955	72%	67%	1.69%

بشكل مشابه لأسهم سوق دمشق، تم تحديد تاريخ بداية التداول الأسهم من 2006/06/05 (5169 يوم) بسبب وجود شركة فينيكس العربية القابضة ذات فترة التداول الأقل. وكان أعلى مدى للبيانات 6955 يوم لكل من شركتي الموارد الصناعية الأردنية وحديد الأردن.

عموماً، أظهرت نتائج المؤشرات السابقة أن أسعار أسهم بورصة عمّان تتميز بحركة مستمرة، الامر الذي يساعد بإعطاء نتائج أكثر فعالية من نموذج التعلم المعزّز العميق حيث تم اختيار الأسهم ذات النسب الأعلى إضافة إلى تلك التي تتضمن أكبر عدد من الأيام والتجربة على بيانات التدريب.

4-2-2- فترات البيانات

فيما يلي تواريخ مجموعات البيانات الفرعية الثلاث التي جزئت اليها البيانات الرئيسية لكل سوق على حدا:

الجدول 3 – مجموعات بيانات الدراسة لكل سوق مالي

Market	Training Set	Validation Set	Test Set
Damascus Securities Exchange	2011/01/09→2018/10/02 2823 days	2018/10/03→2019/09/18 350 days	2019/09/19→2020/07/30 315 days
Amman Stock Exchange	2006/06/05→2017/10/05 4140 days	2017/10/06→2019/03/01 511 days	2019/03/02→2020/07/30 516 days

تستخدم كل من مجموعتي بيانات التدريب والتحقق في نموذج التعلم المعزّز العميق، أما بقية الخوارزميات (الزخم والمعاكسة) فلا تحتاج إلى بيانات تدريب وتحقق وسيتم تطبيقها على بيانات الاختبار مباشرة.

من خلال التجربة، تم الوصول إلى ان حجم نافذة 120 يوم هو الأكثر ملاءمة لهذه البيانات. أي أنه في كل يوم تداول يتم اتخاذ قرار توزيع ثروة المحفظة من خلال استخدام بيانات الأسعار للأسهم المتاحة في الأشهر الأربعة السابقة.

4-2-3- معايير تقييم المحافظ

استخدمت الدراسات السابقة عدة معايير لتقييم النماذج والخوارزميات المستخدمة في التداول. تم انتقاء كل من المعايير التالية لتقييم نتائج التداول لمحافظ البحث:

1) العائد السنوي للمحفظة Annual Return:

هو متوسط العائد اليومي للمحفظة، يحول إلى عائد سنوي من خلال معادلة الفائدة المركبة (Alhabeeb, 2012, 96):

$$(1 + \bar{r})^{250} - 1$$

2) الانحراف المعياري السنوي لعوائد المحفظة Annual Standard Deviation:

يحسب من خلال الانحراف المعياري للعوائد اليومية ومن ثم يتم تحويله إلى سنوي من خلال (Hull, 2015. P.203):

$$STD(r_t) * \sqrt{250}$$

3) القيمة التراكمية للمحفظة Accumulated Portfolio Value:

تعبر عن إجمالي قيمة الاستثمار بعد الانتهاء من عملية التداول. وهي قيمة المحفظة في آخر تاريخ تداول (المعادلة 1، ص.26) (Jiang et al., 2017)

$$\boldsymbol{p_T = p_0 \prod_{t=1}^{T} r_t \cdot w_t}$$

4) نسبة شارب Sharpe Ratio:

هي قيمة علاوة العائد (العائد الذي يزيد عن العائد الخالي من المخاطر) لكل وحدة خطر، وبالتالي يقوم هذا المعيار بالدمج بين مفهوم العائد والخطر، ويحسب من خلال قسمة متوسط علاوة العائد على الانحراف المعياري للعوائد (Cuschieri, 2021, P.10).

$$\boldsymbol{SR = \frac{\bar{r}_- \bar{r}_f}{STD(r_t)}}$$

تم افتراض انعدام معدل العائد الخالي من المخاطرة في كل من سوق دمشق وبورصة عمّان.

5) نسبة الخسارة العظمى Maximum Drawdown:

هي مقياس يعبر عن الانخفاض في قيمة محفظة من أعلى قيمة إلى أقل قيمة وذلك خلال كامل عمرها أو خلال فترة محددة. يحسب من خلال المعادلة التالية (Cuschieri, 2021, P.10):

$$MDD = \frac{Trough\ Value - Peak\ Value}{Peak\ Value}$$

4-2-4- نتائج النماذج المطبقة على سوق دمشق

بسبب عدم توافر طريقة لإدخال مصاريف التداول Transaction Cost ضمن خوارزميات الزخم والمعاكسة سيتم مقارنة نتائجها مع نتائج نموذج التعلم العميق من دون مصاريف تداول (DRL without TC)، لكن سيتم تضمين نموذج التعلم العميق مع مصاريف التداول (DRL-TC) ضمن النتائج لأخذ فكرة عن أدائه في الواقع العملي.

تتراوح نسبة مصاريف التداول المستخدمة في سوق دمشق بين 0.005 و0.008 من قيمة السوقية للأسهم المتداولة[5]. تم استخدام الحد الأعلى (0.008) في نموذج التعلم العميق لمحاكاة أسوأ الأحوال الممكنة.

وكانت النتائج على الشكل التالي:

الجدول 4 - نتائج خوارزميات التداول في سوق دمشق للأوراق المالية على بيانات الاختبار

Metric	Deep Reinforcement Learning		Momentum Strategy Algorithms			Contrarian Strategy Algorithms		
	DRL-TC	**DRL without TC**	**UP**	**EG**	**ONS**	**PAMR**	**OLMAR**	**RMR**
Annual Returns	58.07%	**156.30%**	81.44%	80.63%	49.06%	-61.08%	-59.20%	-55.36%
STD of Returns	21.12%	22.93%	13.49%	13.43%	**10.87%**	23.76%	26.54%	25.40%
APV	1.40	**2.04**	1.58	1.57	1.36	0.47	0.49	0.52
SR	0.14	0.26	**0.280**	0.279	0.23	-0.25	-0.21	-0.20
Max Drawdown	12.66%	6.33%	6.46%	6.45%	**5.71%**	52.96%	52.01%	47.98%

[5] https://web.archive.org/web/20210811112008/http://www.dse.sy/laws-and-regulations/view/23, retrieved on August 15, 2021

يتبين لنا من خلال الجدول (4) أن محفظة التعلم المعزّز العميق DRL without TC قد أعطت أفضل العوائد (156.3%) وأعلى قيمة للمحفظة النهائية (2.04). حيث تشير تلك الأخيرة انه في نهاية السنة الأولى من التداول قد تضاعفت قيمة المحفظة تقريباً.

بشكل عام كانت عوائد خوارزميات الزخم أقل من نموذج التعلم المعزّز العميق DRL without TC، وأعلى من خوارزميات المعاكسة، فقد حققت خوارزمية الزخم UP ثاني أعلى عائد سنوي (81.44%) مقارنة مع جميع الخوارزميات الأخرى. ونلاحظ أن خوارزميات المعاكسة قد حققت أسوأ أداء من حيث العائد السنوي الذي كان سالباً لكل خوارزمية.

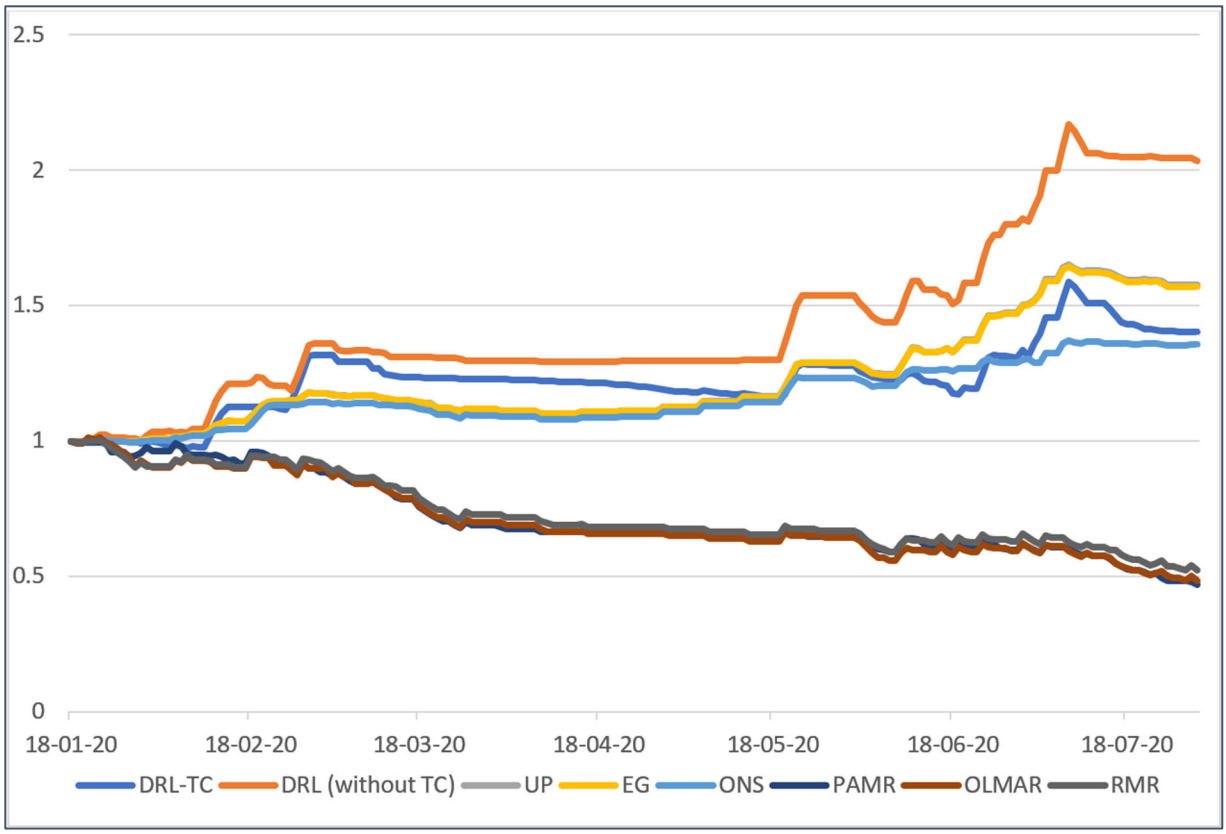
كانت خوارزمية ONS الأكثر أماناً حيث حققت أقل انحراف معياري بما يقارب نصف الانحراف المعياري لنموذج التعلم المعزّز العميق (10.87 % مقارنة مع 22.93%). وكانت نسبة الخسارة العظمى أيضاً منخفضة أي أن أكبر انخفاض في قيمة المحفظة (من أعلى نقطة لقيمة المحفظة إلى أقل نقطة) لم يتجاوز 5.71%.

بالنسبة لنموذج DRL without TC فقد حقق ثاني أفضل نسبة خسارة عظمى (6.33%) وكانت خوارزميات استراتيجية المعاكسة الأسوأ من حيث الخطر أيضا بسبب ارتفاع كل من الانحراف المعياري (أكثر من 23%) ونسبة الخسارة العظمى (أكثر من 47%).

من خلال مقارنة جميع الخوارزميات باستخدام نسبة شارب التي تعكس العائد والخطر معاً، كان ترتيب محافظ خوارزميات الزخم (بالمتوسط) الأعلى، يليه محفظة DRL without TC، ومن ثم محافظ خوارزميات المعاكسة. من الجدير بالذكر أن نسبة شارب تعد كافية للمقارنة بين أداء عدد من المحافظ، لكن من المفيد النظر إلى مكونات هذه النسبة والتحقق من سبب انخفاضها او ارتفاعها، ففي حالة محفظة خوارزمية UP، كان الأداء الأفضل بسبب ارتفاع متوسط العائد السنوي بشكل كبير (ثاني أعلى عائد سنوي) وفي نفس الوقت حققت المحفظة انحراف معياري منخفض نسبياً (ثالث أقل انحراف معياري).

تعد النتائج السابقة أرقام نظرية، أي انه عند التداول الفعلي ستكون النتائج الفعلية أقل وسطيا بسبب وجود مصـــاريف التداول. وهذا ما أظهرته عوائد نموذج DRL–TC فقد كانت أقل بثلاثة أضـــعاف من عوائد النموذج الأســـاســـي (DRL without TC). وكان الانحراف المعياري للعوائد مشـــابها لنموذج DRL without TC. أما نسـبة الخسـارة العظمى كانت ضـعف نموذج DRL without TC. ووصـلت قيمة محفظة DRL–TC إلى 1.4 حيث تشابه بذلك قيمة محافظ خوارزميات الزخم.

الشكل 4 - قيمة المحافظ التراكمية للخوارزميات المطبقة في سوق دمشق على بيانات الاختبار



تحليل النتائج

يوضـح الشـكل (5) متوسـط أسـعار أسـهم سـوق دمشـق التي تم اختيارها، وذلك لكل من فترتي التدريب والتحقق من جهة وفترة الاختبار من جهة أخرى. حيث سـنسـتطيع إلى حد ما تفسـير نتائج الخوارزميات السابقة من خلال تحليل التطور الزمني لهذا المتوسط.

الشكل 5 - متوسط أسعار الأسهم التي تم اختيارها من سوق دمشق

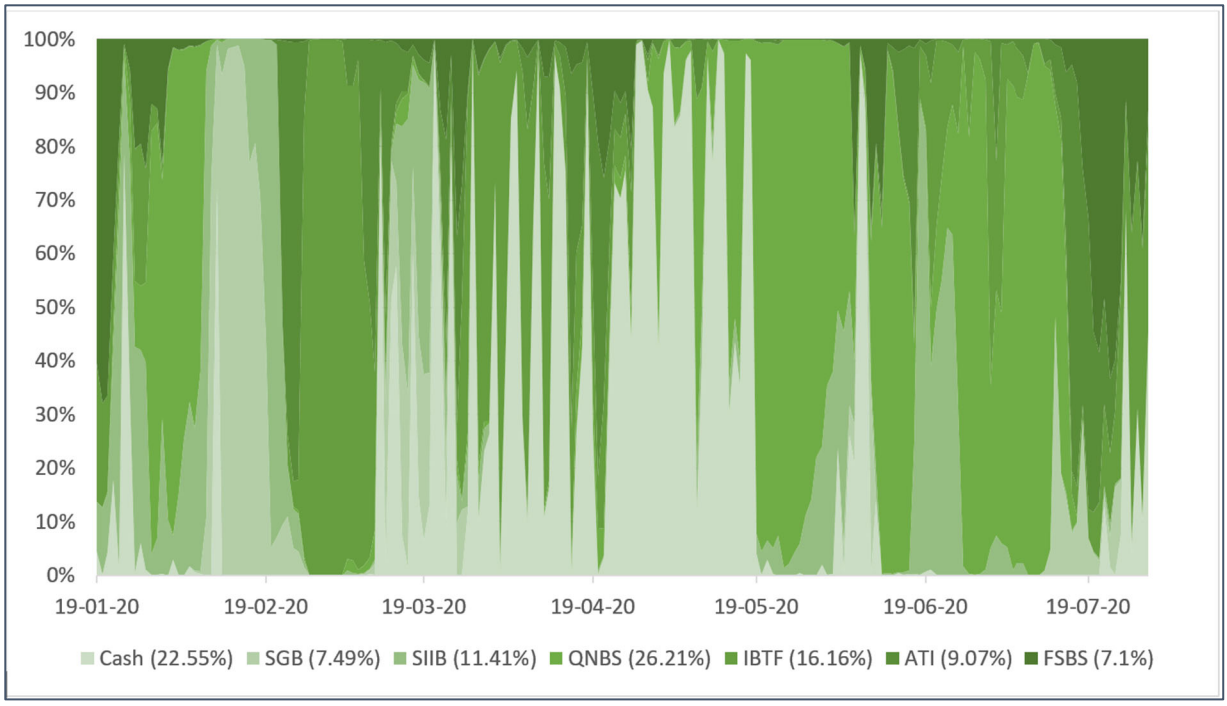
نلاحظ انه خلال فترة التداول كانت أسعار الاسهم وسطيا على ارتفاع مستمر، الأمر الذي يفسر العوائد الإيجابية العالية لخوارزميات الزخم. كما يمكننا تفسير ضعف أداء خوارزميات المعاكسة بنفس السبب، على الرغم من وجود انخفاض في متوسط أسعار الأسهم في الثلث الأول من فترة التداول، لم يكن هذا الأمر كافيا ليؤثر بشكل إيجابي على أداءها.

ويمكننا أيضا تفسير النتائج السابقة من خلال أثر المبالغة في ردود فعل المستثمرين في الأسواق المالية الذي وضّحه (De Bondt and Thaler, 1985)، حيث وجد الباحثان أنه خلال فترات التداول طويلة الأجل يوجد هنالك تغيرات جوهرية في اتجاه عوائد الأسهم. أي أن الأسهم التي أظهرت أقل العوائد خلال السنوات الثلاث إلى الخمسة السابقة ستحقق عوائد جيدة في السنوات الثلاث إلى الخمسة التالية وذلك في سوق الأسهم الأميركي. كما وجدت دراسات مشابهة حدوث الظاهرة نفسها في الأسواق العالمية الأخرى حيث أشارت إلى أن استراتيجيات الزخم تحقق عوائد إيجابية عالية في الفترات قصير الاجل، أي من ستة أشهر إلى حد السنة.

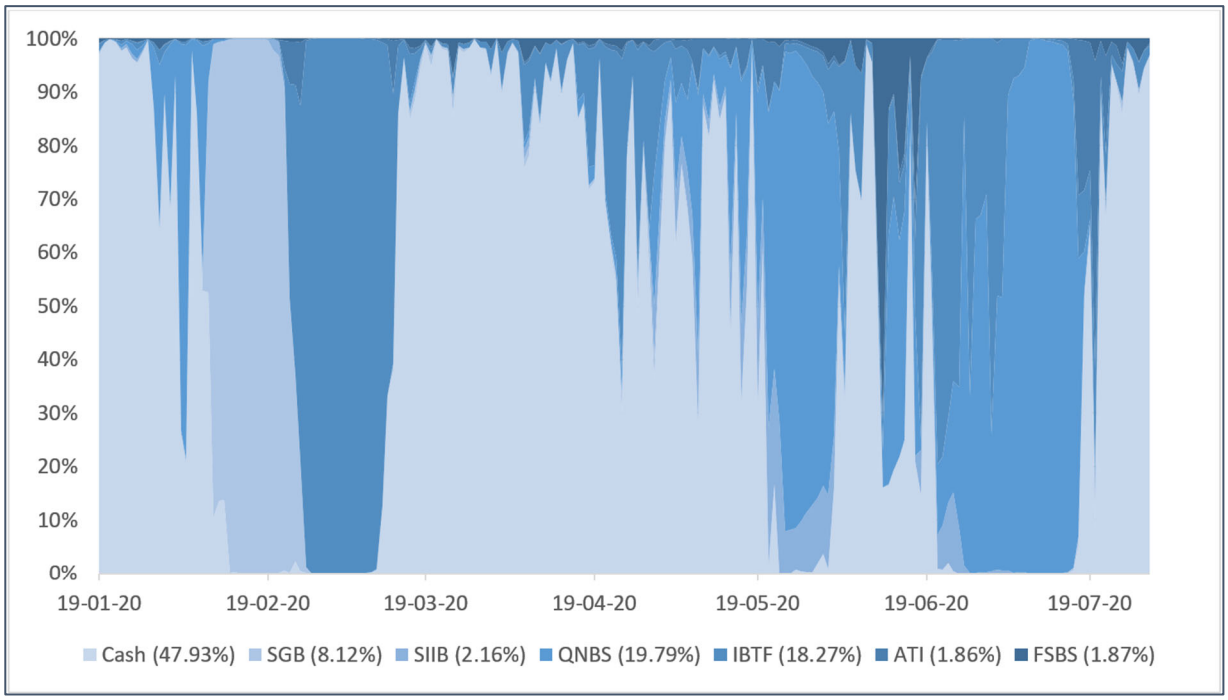
وبالتالي فان مدة التداول القصيرة في سوق دمشق (195 يوماً دون النافذة) لم تكن كافية لحدوث انعكاس السعر الموضح بالظاهرة السابقة والذي يعد ضروري حتى تحقق خوارزميات المعاكسة عوائد إيجابية، لكنها كانت أكثر من كافية بالنسبة لخوارزميات الزخم.

بالنسبة إلى نموذج DRL without TC، كانت عوائده مرتفعة بسبب قدرته على تعلم أنماط حركة كل سهم خلال فترة التدريب. فقد واجه حالتين مشابهتين للارتفاع بقيمة الأسهم الذي حصل في فترة الاختبار. حيث بدأت الحالة الأولى في منتصف الشهر الأول من عام 2017 وكانت الحالة الثانية في بداية الشهر العاشر من نفس السنة كما هو ظاهر في الشكل (5).

الشكل 6 - تركيبة محفظة نموذج DRL without TC في سوق دمشق على بيانات الاختبار



الشكل 7 - تركيبة محفظة نموذج DRL-TC في سوق دمشق على بيانات الاختبار



يظهر الشكل (6) والشكل (7) تركيبة محفظة من الأسهم والنقد خلال فترة التداول لكل من نموذج DRL without TC وDRL–TC على التوالي. نلاحظ بأن نموذج DRL–TC كان أكثر تحفظا ببيعه وشـراءه للأسـهم، حيث احتوت محفظته على النقد أو سـهم واحد لفترات أطول من نموذج DRL without TC. وهو امر منطقي بسبب وجود مصـاريف التداول، والتي تعد السبب الرئيسـي لانخفاض عوائد DRL–TC حيث تسـبب تآكل في قيمة المحفظة على مدى عمرها، كما أدى ارتفاع نسـبة النقد ضـمن ممتلكات المحفظة إلى ضـياع فرص توظيف الأموال وانخفاض العائد، فقد ازداد متوسـط أوزان النقد خلال فترة التداول من 22.55% إلى 47.93% من إجمالي أصول المحفظة.

4-2-5- نتائج النماذج المطبقة على بورصة عمّان

في بورصـــة عمّان، تتراوح مصـــاريف التداول بين 0.0054 و0.0074 من القيمة الســـوقية للأســـهم المتداولة[6]. يضـــاف إلى هذه النســـبة 0.0008 من قيمة الأســـهم الســـوقية كضـــرائب تدفع من كل طرف لخزينة الدولة. فاذا اعتمدنا الحد الأعلى كما في حال بيانات سوق دمشق تصبح إجمالي مصاريف التداول 0.0074 + 0.0008 = 0.0082.

الجدول 5 - نتائج خوارزميات التداول في بورصة عمّان على بيانات الاختبار

Metric	Deep Reinforcement Learning		Momentum Strategy Algorithms			Contrarian Strategy Algorithms		
	DRL-TC	**DRL without TC**	**UP**	**EG**	**ONS**	**PAMR**	**OLMAR**	**RMR**
Annual Returns	-0.62%	**165.90%**	-1.78%	-1.80%	0.36%	-29.42%	-23.71%	-31.75%
STD of Returns	30.67%	37.09%	9.46%	**9.43%**	11.08%	35.78%	38.12%	37.93%
APV	0.92	**4.23**	0.97	0.96	0.996	0.52	0.58	0.49
SR	-0.001	**0.17**	-0.01	-0.01	0.002	-0.06	-0.04	-0.06
Max Drawdown	40.93%	34.03%	18.29%	18.26%	**16.66%**	57.52%	65.17%	66.34%

نلاحظ من الجدول (5) تفوق محفظة DRL without TC على جميع المحافظ الأخرى بشـكل هائل من حيث العوائد (165.9%)، وقيمة المحفظة التراكمية التي وصـلت إلى 4.23 ضـعف قيمة الاسـتثمار في نهاية فترة التداول.

بالمقابل، جميع محافظ الخوارزميات الأخرى (زخم ومناقضـــة) حققت عوائد ســـلبية، عدا خوارزمية ONS (0.36%)، إلا أن قيمة محفظتها في نهاية فترة التداول كانت 0.996 أي خســـرت ما يعادل 0.004

[6] https://web.archive.org/web/20210301044839/https://www.ase.com.jo/en/Product-Services/Trading-Services/Transaction-Cost, retrieved on August 16, 2021

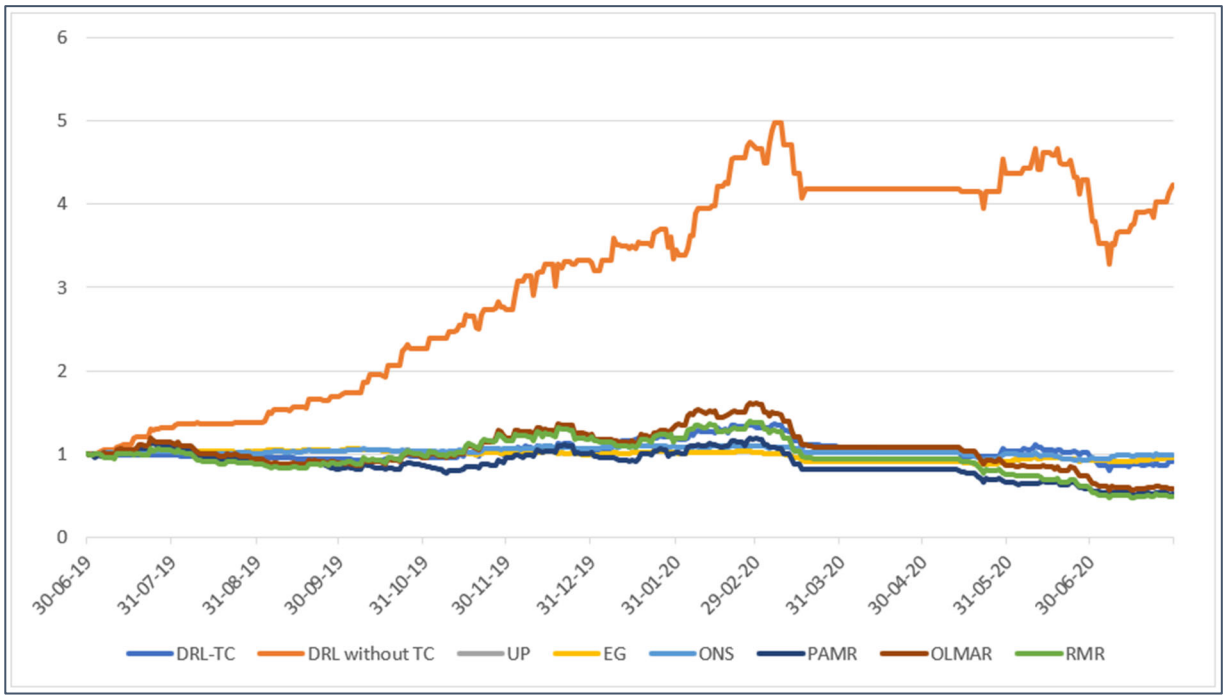
دينار عن كل دينار مستثمر. كانت العوائد السلبية لمحافظ خوارزميات المعاكسة وسطيا أكبر بمعدل 26 مرة من العوائد السلبية لمحافظ خوارزميات الزخم.

من خلال النظر إلى معايير الخطر نجد أن خوارزمية EG هي الأفضل، فقد وصل الانحراف المعياري إلى 9.43%، أي ما يعادل ربع الانحراف المعياري لخوارزمية DRL without TC (37.09%)، وحققت خوارزمية ONS أقل نسبة خسارة عظمى (16.66%) يليها خوارزمية EG (18.26%).

تعكس هذه النتائج نسبة شارب، حيث انه بالمتوسط كل 1% من المخاطر أو التقلبات المقاسة بالانحراف المعياري التي تتحملها محفظة التعلم المعزّز العميق DRL without TC يقابلها 0.17% من العائد، وأن كل 1% من مخاطر خوارزمية ONS تحقق ما يعادل 0.002% من العوائد. إن نسبة شارب لمحفظة DRL without TC كانت أعلى بـ 85 مرة من نسبة شارب لخوارزمية ONS.

بالنسبة لنموذج DRL–TC فقد حقق خسائر سنوية بمعدل 0.62%، حيث انخفضت قيمة المحفظة من 1 في نهاية فترة التداول إلى 0.92 (0.08 دينار)، إضافة إلى أن مؤشرات الخطر كانت مرتفعة (30.67% و40.93% لكل من الانحراف المعياري ونسبة الخسارة العظمى على التوالي)، حيث أدى ذلك إلى انخفاض نسبة شارب بشكل كبير (عدا عن كونها سالبة بسبب العوائد السالبة).

الشكل 8 - قيمة المحافظ التراكمية للخوارزميات المطبقة في بورصة عمّان على بيانات الاختبار



تحليل النتائج

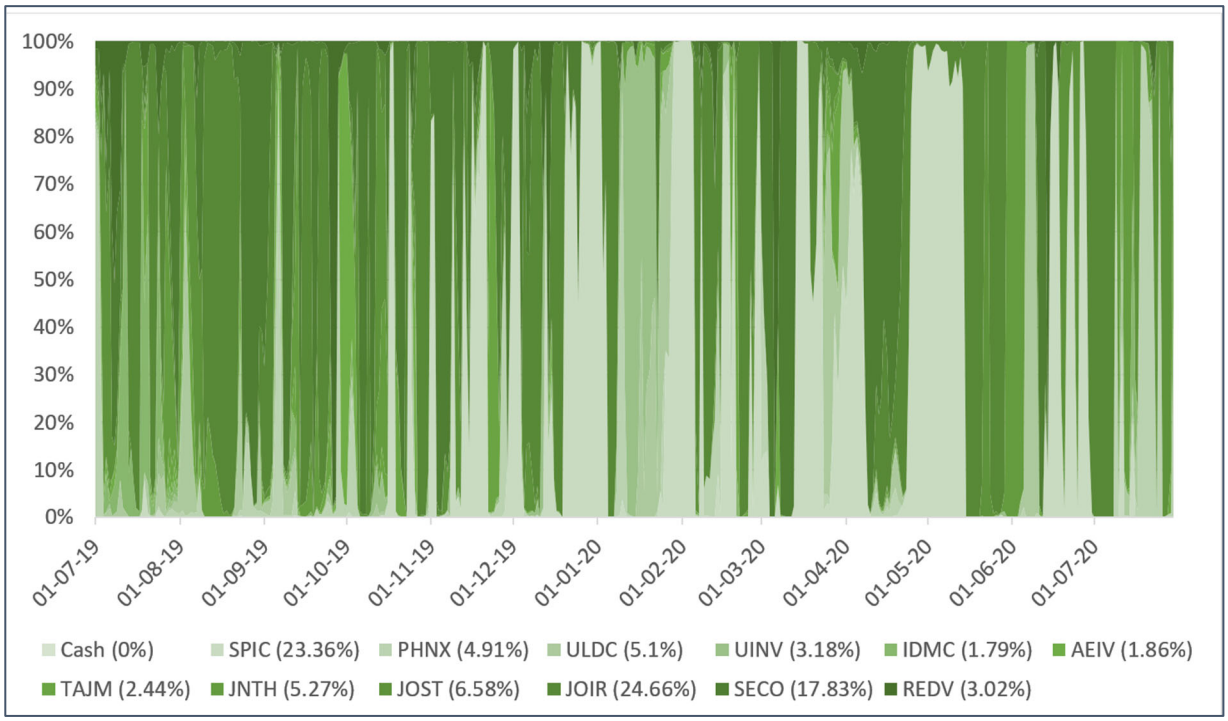
الشكل 9 - متوسط أسعار الأسهم التي تم اختيارها من بورصة عمّان

يتبين من الشـــكل (9) انه لم يكن هنالك تغير جوهري في متوســـط أســـعار الأســـهم خلال فترة التداول (الاختبار). حيث أنه قد ازداد قليلاً في الفترة ما بين بداية شـــهر 7 عام 2019 ومنتصـــف شـــهر 3 عام 2020. إلا أنه عاود الانخفاض فيما بعد وبقي على نفس المســـتوى حتى نهاية فترة التداول التي عادت فيها قيمته إلى ما كانت عليه في البداية (من 0.57 دينار إلى 0.5708 دينار).

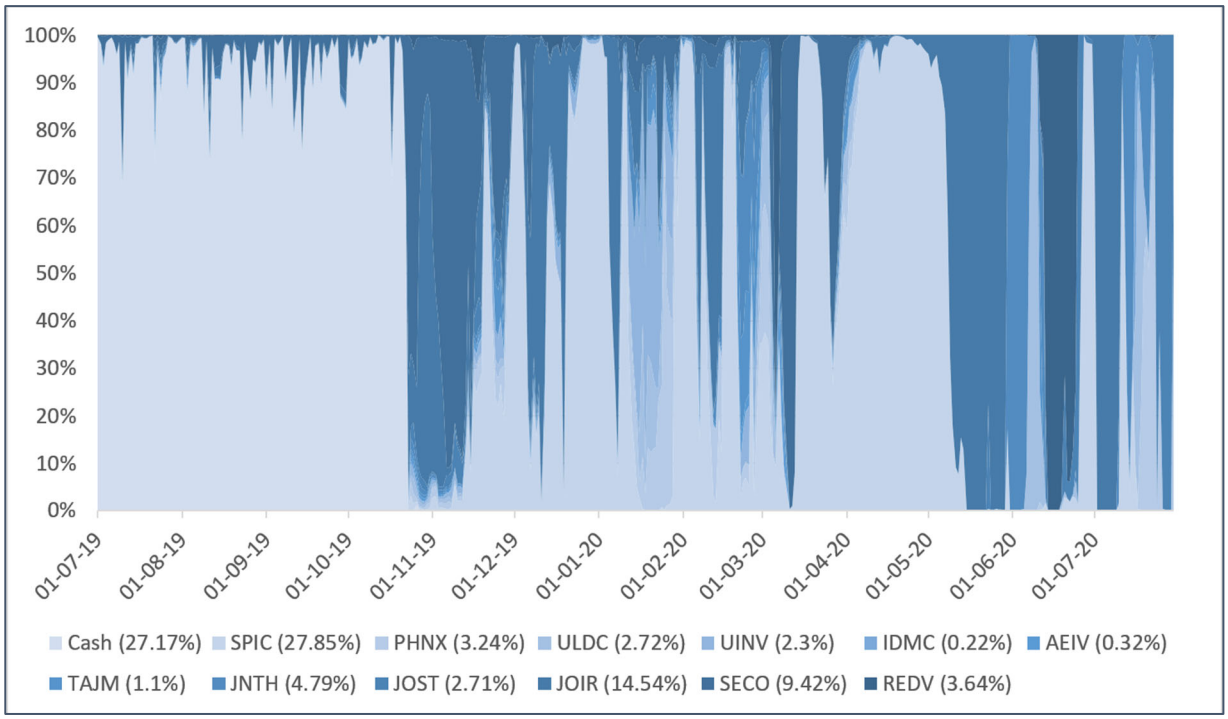
على الرغم من أن فترة التداول كانت بحدود السنة (396 يوم)، أي ملائمة لتعطي خوارزميات الزخم عوائد إيجابية، إلا أنه بسـبب ثبات الأسـعار وسـطياً خلال الفترة، لم تكن تلك الخوارزميات قادرة على تحقيق أي عائد. أما بالنسبة لخوارزميات المعاكسة فان سوء عوائدها يفَسر بانخفاض فترة التداول (فهي تصلح لفترات التداول طويلة الأجل).

يعتمد نموذجي DRL وDRL−TC على التقلبات قصـــيرة الاجل لتحقيق العوائد، لكن كما هو موضـــح بالشكل (9) هذه التقلبات لم تكن شديدة ومتكررة بشكل يكفي لتعويض مصاريف التداول، الامر الذي جعل نسبة الخسارة العظمى المرتفعة لكل منهما وأيضاً يفسر الخسائر المحققة من قبل نموذج DRL−TC.

الشكل 10 - تركيبة محفظة نموذج DRL without TC في بورصة عمّان على بيانات الاختبار



الشكل 11 - تركيبة محفظة نموذج DRL-TC في بورصة عمّان على بيانات الاختبار



نلاحظ من الشـكلين (10) و (11) وجود نفس الأثر لمصـاريف التداول في تركيبة المحفظة لتلك المبنية من أسـهم سـوق دمشـق، فقد ازدادت نسـبة متوسـط النقد في إجمالي أصـول المحفظة من 0% إلى 27.17% في نموذج DRL–TC، مما أدى إلى انخفاض توظيف الأموال، وتآكل قيمة المحفظة نتيجة مصاريف التداول، الامر الذي أدى إلى العوائد السالبة.

4-3- المبحث الثالث – اختبار الفرضيات والنتائج والتوصيات

تمهيد

يسعى هذا المبحث إلى الإجابة عن أسئلة البحث من خلال اختبار فرضيات البحث لكل سوق على حدا. ومن ثم مناقشة النتائج التي توصل اليها وعرض التوصيات المبنية على هذه النتائج. وفي النهاية سيتم عرض المجالات المستقبلية للبحث.

4-3-1- اختبار الفرضيات

فرضية البحث الأساسية: يوجد اختلاف جوهري في الاداء بين خوارزميات الإدارة المستمرة لمحافظ الاسهم

تستخدم نسبة شارب لتقييم أداء محفظة مالية. حيث تجمع بين مفهومي العوائد والمخاطر في قيمة واحدة. لكن هذه النسبة لوحدها لا تحمل دلالة أو معنى، حيث أن الاستخدام الأساسي لها هو للمقارنة بين محفظتين أو أكثر.

للوصول إلى نتيجة ذات فائدة وقابلة للانعكاس على أرض الواقع، يجب معرفة ما إذا كانت الفروقات الناتجة بين النسبتين ذات دلالة إحصائية، وأن الفرق لم يكن نتيجة التذبذبات الحاصلة في أسعار السهم ضمن العينة المتوفرة. أي أنه نريد معرفة ما إذا كان هنالك فرق حقيقي في الأداء بين الاصلين، ويمكن استخدامه لبناء القرارات الاستثمارية.

وبالتالي يجب تطبيق إحدى الاختبارات الإحصائية للوصول إلى هذه النتيجة. تم استخدام اختبار t-test المعدل وفق (Opdyke, 2007) الذي لا يشترط أن تكون عوائد المحفظة موزعة طبيعياً على المتغير العشوائي الناتج عن طرح نسبة شارب لمحفظتين ماليتين.

تصاغ فرضية العدم والفرضية البديلة التي سيتم اختبارهم على الشكل التالي:

$$H_0: (SR_B - SR_A) \leq 0$$
$$H_1: (SR_B - SR_A) > 0$$

حيث أن SR_B: هي نسبة شارب للمحفظة الثانية

و SR_A: هي نسبة شارب للمحفظة الأولى

بهدف اختبار الفرضيات على استراتيجيات الزخم والمعاكسة ككل. تم تجميع نتائج خوارزميات الزخم في محفظة واحدة عن طريق حساب متوسط العوائد اليومية لخوارزميات الزخم الثلاثة (UP, EG وONS). وبنفس الأمر تمت معالجة خوارزميات المعاكسة للحصول على محفظة واحدة تمثل الخوارزميات الثلاث (PAMR, OLMAR وRMR). حيث أصبح لدينا سلسلة من عوائد محفظة استراتيجية الزخم وسلسلة من عوائد محفظة استراتيجية المعاكسة نستطيع استخدامها في حساب نسبة شارب والتحقق من معنوية الفروقات بين كل زوج.

كانت قيمة نسبة شارب لمحفظة استراتيجية الزخم ومحفظة استراتيجية المعاكسة لكل من سوقي دمشق وعمّان كما يلي:

الجدول 6 - نسبة شارب لكل نوع من المحافظ الثلاث في كل من سوق دمشق وبورصة عمّان

Market	DRL-TC	DRL (without TC)	Momentum	Contrarian
DSE	0.14	0.26	0.27	-0.23
ASE	0.00	0.17	-0.01	-0.06

يعرض الجدول (7) نتائج اختبار الفروقات الاحصائي حيث يتضمن قيمة فرق نسبة شارب ونسبة p− value لكل زوج من المحافظ المتاحة في سوقي دمشق وعمّان. على الشكل التالي:

الجدول 7 - نتائج اختبار الفروقات الاحصائي لنسب شارب في كل من سوق دمشق وسوق عمّان

Damascus Stock Exchange						
A/B	DRL without TC		Momentum		Contrarian	
	SR Diff.	P-Value	SR Diff.	P-Value	SR Diff.	P-Value
DRL-TC	0.12	0.3378%	0.13	1.13%	-0.37	100%
DRL without TC			0.01	42%	-0.49	100%
Momentum					-0.51	100%
Amman Stock Exchange						
A/B	DRL without TC		Momentum		Contrarian	
	SR Diff.	P-Value	SR Diff.	P-Value	SR Diff.	P-Value
DRL-TC	0.17	0.0006%	-0.01	54.96%	-0.06	87%
DRL without TC			-0.17	100%	-0.23	100%
Momentum					-0.05	82%

فرضـية البحث الفرعية (1): يوجد اختلاف في أداء المحافظ وفق نسـبة شـارب وذلك لصـالح محفظة التعلم المعزّز العميق مقارنة مع محفظة استراتيجية الزخم.

$$H_0: (SR_{Momentum} - SR_{DRL\ without\ TC}) \leq 0$$

$$H_1: (SR_{Momentum} - SR_{DRL\ without\ TC}) > 0$$

من خلال النظر إلى الجدول (7) لسـوق دمشـق، نلاحظ بأن الفرق بين نسـبة شـارب لمحفظة DRL without TC ونسبة شارب لمحفظة الزخم هو 0.01 (لصالح محفظة الزخم)، لكن p−value هي 42% وهي أكبر من أي حد ثقة مسـتخدم وبالتالي سـيتم قبول فرضـية العدم القائلة بأن نسـبة شـارب لمحفظة DRL without TC هي أكبر من نسبة شارب لمحفظة استراتيجية الزخم، أي ان الفرق الظاهر هو ليس معنوي احصائياً وسيتم رفض الفرضية البديلة.

الامر مماثل بالنسبة لبورصـة عمّان فقد قبلنا فرضـية العدم ورفضـنا الفرضـية البديلة لأن قيمة p−value تعادل 100% أي أنه لا يوجد دلالة إحصـائية بان نسـبة شـارب لمحفظة اسـتراتيجية الزخم هي أكبر من نسبة شارب لمحفظة DRL without TC في بورصة عمّان.

نتائج الفرضية الفرعية (1): قبول لكل من سوق دمشق وبورصة عمّان

فرضـية البحث الفرعية (2): يوجد اختلاف في أداء المحافظ وفق نسـبة شـارب وذلك لصـالح محفظة التعلم المعزّز العميق مقارنة مع محفظة استراتيجية المعاكسة.

$$H_0: (SR_{Contrarian} - SR_{DRL\ without\ TC}) \leq 0$$

$$H_1: (SR_{Contrarian} - SR_{DRL\ without\ TC}) > 0$$

يشير الجدول (6) إلى أن محفظة المعاكسة كانت ذات أداء سيئ بسبب قيم نسبة شارب السالبة التي كانت نتيجة العوائد السالبة (الخسائر) التي حققتها خوارزميات المعاكسة.

بالمقارنة مع محفظة التعلم المعزّز العميق (الجدول (7))، نجد أن p−value في سوق دمشق هي 100% أي أنه سيتم قبول فرضية العدم التي تشير إلى وجود فرق معنوي في نسبة شارب لصالح محفظة التعلم المعزّز العميق.

كما نرى الأمر مماثل في بورصة عمّان، فقد كانت ال p−value أيضا 100%. أي ان الفرق الناتج بين النسبتين لم يكن بسبب التذبذبات الحاصلة في أسعار الأصول وهو فرق فعلي متأصل في طبيعة الخوارزميات المستخدمة.

نتائج الفرضية الفرعية (2): قبول لكل من سوق دمشق وبورصة عمّان

فرضـية البحث الفرعية (3): يوجد اختلاف في أداء المحافظ وفق نسـبة شـارب وذلك لصـالح محافظ استراتيجية الزخم مقارنة مع محفظة استراتيجية المعاكسة.

$$H_0: (SR_{Contrarian} - SR_{Momentum}) \leq 0$$

$$H_1: (SR_{Contrarian} - SR_{Momentum}) > 0$$

بمقارنة نسبة شارب لمحفظة استراتيجية الزخم مع محفظة استراتيجية المعاكسة نجد التفوق الظاهر لمحفظة الزخم، حيث كان الفرق هو الأكبر بين الاثنين (في سوق دمشق 0.51) بسبب العوائد السالبة لمحفظة المعاكسة. وهو فرق معنوي احصائياً حيث كانت p−value 100% وبالتالي سنقبل فرضية العدم التي تشير إلى وجود فرق معنوي بين قيمتي نسبة شارب لصالح محفظة الزخم في سوق دمشق.

كانت النتائج مماثلة في بورصـــة عمّان، إلا أن الفرق لم يكن بنفس المســتوى (0.05) لكنه أيضــا معنوي احصائياً لان p–value هي 82% أي انه سيتم رفض الفرضية البديلة وقبول فرضية العدم التي تقول بأن نسبة شارب لمحفظة الزخم هي أكبر من نسبة شارب لمحفظة المعاكسة في بورصة عمّان.

نتائج الفرضية الفرعية (3): قبول لكل من سوق دمشق وبورصة عمّان

وبالتالي سيتم قبول فرضـية البحث الأسـاسـية القائلة بأنه يوجد اختلاف جوهري في الأداء بين خوارزميات الإدارة المستمرة لمحافظ الأسهم، وذلك بسبب قبول الفرضيات الفرعية بالكامل لوجود دليل احصائي معنوي على ذلك.

4-3-2- النتائج

توصل التطبيق العملي لخوارزميات إدارة المحفظة المستمرة على سوقي دمشق وعمّان إلى النتائج التالية:

1- يمكن ملاحظة التفوق الظاهر لمحفظة التعلم المعزّز العميق على محفظتي الزخم والمعاكســـة في كل من ســوق دمشــق وبورصــة عمّان، ووجد ان الفرق معنوي إحصــائياً عند اختبار الفرق بين قيم نســبة شـــارب وهو يتفق مع نتائج (Jiang et al., 2017)، (Guo et al., 2018)، (Gao et al., 2020) و (Zhang et al., 2020)

2- وُجد فرق بين محفظتي الزخم والمعاكسة لصـالح محفظة الزخم وتم تفسيره في سوق دمشق من خلال المؤشــر الصــاعد لمتوســط أســهم الســوق، فهو دليل على قدرة الخوارزميات على اســتغلال الظاهرة وتحقيق عوائد مرتفعة. وأن ســبب العوائد الســلبية لمحافظ المعاكســة هو فترة التداول القصــيرة التي لم تكن كافية لظهور أثر المعاكســـة في أســـعار الأصـــول. فان فترة الاحتفاظ يجب أن تصـــل إلى حد الخمس ســنوات على الأقل حتى تســتطيع اســتراتيجية المعاكســة تحقيق عوائد معنوية (De Bondt and Thaler, 1985). وهــذه النتــائج متوافقــة مع دراســـتي (Jiang et al., 2017) و (Guo et al., 2018).

3- نلاحظ تشـــابه الأداء إلى حد ما بين خوارزميات الزخم من حيث العوائد والخطر، وذلك لأنها مبنية على نفس الأثر من جهة وبسـبب تشـابه التطبيق البرمجي لها من جهة أخرى. والأمر شـبيه بالنسـبة لخوارزميات المعاكسة.

4- كما لوحظ اختلاف في أداء خوارزمية ONS تحديداً، حيث كانت الخوارزمية الأكثر اماناً في ســـوق دمشــق من حيث الانحراف المعياري للعوائد ونســبة الخســارة العظمى، متفوقة بذلك على خوارزميتي UP وEG. أما في بورصة عمّان فكانت ذات الانحراف المعياري الأعلى ونسبة الخسارة الأقل مقارنة مع خوارزميات الزخم الأخرى. الأمر الذي يمكن تفسيره باختلاف اتجاه منحنى متوسط أسعار الأسهم في السوقين.

4-3-3- التوصيات

بناءً على ما تم التوصل اليه من نتائج في الدراسة العملية، توصي الدراسة المستثمرين بما يلي في سوق دمشق للأوراق المالية:

1- بشكل عام ينصح باستخدام نموذج التعلم المعزّز العميق DRL-TC بسبب عوائده المرتفعة من جهة ولأنه يراعي مصاريف التداول من جهة أخرى، أي أن نتائجه العملية ستكون قريبة من النتائج التي تم التوصل اليها في هذه الدراسة. لكن يمكن أن تكون مؤشرات الخطر المرتفعة لهذا النموذج غير محفزاً بالنسبة للمستثمرين غير المحبين للمخاطر.

2- استخدام خوارزمية ONS للمستثمرين الذين لا يحبذون المخاطرة، فهي الأكثر أمانا بسبب انخفاض مؤشرات الخطر المنخفضة لها مقارنة مع الخوارزميات الاخرى.

3- استخدام خوارزميات الزخم بشكل عام (وتجنب خوارزميات المعاكسة) عند الاستثمار قصير الاجل في سوق دمشق.

وتوصي الدراسة المستثمرين بما يلي في بورصة عمّان:

1- ينصح باستخدام نموذج التعلم المعزّز العميق DRL-TC في بورصة عمّان على الرغم من العوائد الشبه معدومة. حيث يمكن أن تصبح إيجابية ومرتفعة في حال كان هنالك تذبذب في أسعار الأسهم. وأن مراعاته لمصاريف التداول تعطيه الأفضلية على بقية الخيارات. وأنه في هذه الدراسة تم اعتماد أعلى مصاريف تداول متاحة، حيث أنه في الواقع العملي لا يتم دفع أعلى مبلغ للمصاريف في جميع الأحوال وبالتالي فانه من المعقول الافتراض أنه ستكون العوائد عالية عند الاستثمار الفعلي.
2- الابتعاد عن خوارزميات المعاكسة بسبب خسائرها المرتفعة في الأجل القصير.
3- استخدام خوارزمية ONS من خوارزميات الزخم لكن بحذر.

4-3-4- افاق البحث المستقبلية

تقترح الدراسة التوسع في البحث وإجراء دراسات في المجالات التالية:

1- إيجاد طريقة لإدخال مصاريف التداول في خوارزميات الزخم والمعاكسة، بهدف الحصول على نتائج أكثر واقعية وقابلية للتطبيق في الواقع العملي.
2- تنويع فترات الدراسة عند توافر كمية بيانات مناسبة، وذلك بهدف التحقق من الأداء طويل الأجل لكل من نموذج التعلم المعزّز العميق وخوارزميات الزخم والمعاكسة.
3- اختبار نموذج التعلم المعزّز العميق على الأسواق العربية الأخرى باعتباره حديث نسبياً وله مستقبل واعد.
4- بالإضافة إلى الأسعار التاريخية المستخدمة في هذا البحث، يمكن استخدام البيانات السلوكية كالأخبار المالية النصية والمعلومات المستخرجة من وسائل التواصل الاجتماعي في نموذج تعلم معزّز عميق ومقارنة أداءه مع الخوارزميات الأخرى.

قائمة المراجع

1- المراجع العربية

الكتب

1. موصلي سليمان، سليمان عدنان، الاسواق المالية، جامعة دمشق، 2013
2. الزرقان صالح طاهر، العوامل المالية والاقتصادية المؤثرة في عوائد الأسهم: النظرية والتطبيق، المنهل، 2010
3. ربابعة عبد الرؤوف، بناء المحافظ الاستثمارية وإدارة الاستثمار في الأسهم بين العوائد والمخاطرة، هيئة الأوراق المالية، الاردن، 2009
4. كنعان علي، الأسواق المالية، مطبعة الروضة، سورية، 2009
5. هوشيار معروف، الاستثمارات والأسواق المالية، دار صفاء للنشر والتوزيع، الاردن، 2009
6. عبد الجواد محمد عوض، الشديفات علي إبراهيم، الاستثمار في البورصة أسهم - سندات - أوراق مالية، دار الحامد للنشر والتوزيع، الأردن، 2006
7. العتيبي أحمد معجب، المحافظ المالية الاستثمارية - أحكامها وضوابطها في الفقه الإسلامي، دار النفائس للنشر والتوزيع، الاردن، 2005
8. الهندي منير إبراهيم، ادوات الاستثمار في اسواق راس المال: الاوراق المالية وصناديق الاستثمار، المكتب العربي الحديث، مصر، 2003
9. خريوش حسني، ارشيد عبد المعطي رضا، جودة محفوظ احمد، إدارة المحافظ الاستثمارية، دار زهران للنشر والتوزيع، الاردن، 1999

أبحاث غير منشورة

1. بوكرومة كريمة، أثر قرارات الاستثمار وتوزيع الأرباح على الأداء المالي للشركات الصناعية المدرجة في سوق عمّان للأوراق المالية للفترة (2000 - 2018)، أطروحة دكتوراه، جامعة العربي بن مهيدي، 2019
2. قبنض رشا، دور المعرفة المالية في مساعدة المؤسسات على بناء نموذج لإدارة المحافظ الاستثمارية، أطروحة دكتوراه، جامعة حلب، 2018
3. باخوص روى، دور هيئة الأوراق والأسواق المالية السورية في رفع كفاءة السوق المالي (دراسة تطبيقية على سوق دمشق للأوراق المالية)، رسالة ماجستير، جامعة دمشق، 2015

2- المراجع الأجنبية

Books

1. Mehlig B., Machine Learning with Neural Networks: An Introduction for Scientists and Engineers, Cambridge University Press, United Kingdom, 2021
2. Umbrello S., AI Winter (Encyclopedia of Artificial Intelligence: The Past, Present, and Future of AI), ABC-CLIO, United States, 2021
3. Géron A., Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow Concepts, Tools, and Techniques to Build Intelligent Systems, O'Reilly Media, United States, 2019
4. Beverungen A., Algorithmic Trading, Artificial Intelligence and the Politics of Cognition, Andreas Sudmann: The democratization of artificial intelligence. Net Politics in the Era of Learning Algorithms, 2019, P.77-93
5. Li B. and Hoi S., Online Portfolio Selection: Principles and Algorithms, CRC Press, United States, 2018
6. Sutton R., Barto A., Reinforcement Learning: An Introduction second edition, MIT Press, Massachusetts, United States, 2018
7. Hull J., Risk Management and Financial Institutions 4th Edition, Wiley, United States, 2015
8. Nielsen M., Neural Networks and Deep Learning, http://neuralnetworksanddeeplearning.com, 2015
9. Alpaydin E., Introduction to Machine Learning, MIT Press, United States, 2014
10. Kissell R., The Science of Algorithmic Trading and Portfolio Management, Elsevier Science, 2013
11. Alhabeeb M., Mathematical Finance, Wiley, United States, 2012
12. Spaan M., Partially Observable Markov Decision Processes, Adaptation, Learning, and Optimization, Springer, Gemany, 2012, P.387-414
13. Bäuerle N. and Rieder U., Markov Decision Processes with Applications to Finance, Springer, Germany, 2011
14. Kirkpatrick C. and Dahlquist J., Technical Analysis: The Complete Resource for Financial Market Technician, Pearson Education Incorporated, United Kingdom, 2006
15. Gallati R., Risk Management and Capital Adequacy, McGraw-Hill Education, United States, 2003
16. Harris L., Trading and Exchanges Market Microstructure for Practitioners, Oxford University Press, United Kingdom, 2003
17. Anthony M., Neural Network Learning: Theoretical Foundations, Cambridge University Press, United States, 1999
18. Haykin S., Neural Networks: A Comprehensive Foundation, Prentice Hall, New Jersey, 1999

19. Werbos P., The Roots of Backpropagation: From Ordered Derivatives to Neural Networks and Political Forecasting, Wiley, United States, 1994
20. Thorndike E., Animal Intelligence (series): Experimental Studies, Transaction Publishers, United States, 1970
21. Minsky M., Papert S., Perceptrons an Introduction to Computational Geometry, MIT Press, United States, 1969

Journals and Conferences

1. Huang Z., Tanaka F., MSPM: A modularized and scalable multiagent reinforcement learning-based system for financial portfolio management, PLoS ONE 17(2), 2022
2. Cuschieri N., Vella V., Bajada J., TD3-Based Ensemble Reinforcement Learning for Financial Portfolio Optimisation, 31st International Conference on Automated Planning and Scheduling, 2021, P.6-14
3. Gao Z., Gao Y., Hu Y., Jiang Z., Su J., Application of Deep Q-Network in Portfolio Management, 5th IEEE International Conference on Big Data Analytics, 2020
4. Luo T., Wu M., Chen C., A Framework of Deep Reinforcement Learning for Stock Evaluation Functions, Journal of Intelligent and Fuzzy Systems Vol. 38 (7), 2020, P.1-11
5. Weng L., Sun X., Xia M., Liu J., Xu Y., Portfolio Trading System of Digital Currencies: A Deep Reinforcement Learning with Multidimensional Attention Gating Mechanism, Neurocomputing Vol. 402, 2020, P.171-182
6. Zhang Y., Zhao P., Li B., Huang L., Wu Q., Tan M., Cost-Sensitive Portfolio Selection via Deep Reinforcement Learning, IEEE Transactions on Knowledge and Data Engineering, 2020
7. Pendharkar P., Cusatis P., Trading Financial Indices with Reinforcement Learning Agents, Expert Systems with Applications, United States, V. 103, 2018, P.1-13
8. Pouyanfar S, Sadiq S., Yan Y., Haiman T., Tao Y., Reyes M., Shyu M., Chen S., Iyengar S., A Survey on Deep Learning: Algorithms, Techniques, and Applications, ACM Computing Surveys, V. 51 No. 5, 2018, P.1-36
9. Olah C., Mordvintsev A., Schubert L., Feature Visualization, Distill, 2017
10. Szegedy C., Ioffe S., Vanhoucke V., Alemi A., Inception-v4 - Inception-ResNet and the Impact of Residual Connections on Learning, Thirty-First AAAI Conference on Artificial Intelligence, 2017
11. Li B. and Hoi S., Online Portfolio Selection: A Survey, ACM Computing Surveys Vol. 46 (3), 2014, P.1-36
12. Liston, C., Cohen, M., Teslovich, T., Levenson, D., & Casey, J., Atypical Prefrontal Connectivity in Attention-Deficit/Hyperactivity Disorder: Pathway to Disease or Pathological End Point?, Biological Psychiatry, 2011, P.1168-1177

13. Opdyke J.D., Comparing Sharpe Ratios: So Where are the p-values?, Journal of Asset Management, Vol. 8, 2007, P.308-336
14. Agarwal A., Hazan E., Kale S., Schapire R., Algorithms for Portfolio Management based on the Newton Method, Proceedings of the 23rd International Conference on Machine Learning, Pittsburgh, 2006
15. Lo A., Mamaysky H., Wang J., Foundations of Technical Analysis: Computational Algorithms, Statistical Inference, and Empirical Implementation, The Journal of Finance Vol. 55(4), 2000, P.1705–1770
16. Bertsimas D. and Lo A., Optimal Control of Execution Costs, Journal of Financial Markets, V. 1, 1998, P.1-50
17. Helmbold D., Schapire R., Singer Y., Warmuth M., Online Portfolio Selection using Multiple Updates, Mathematical Finance Vol. 8 (4), 1998, P.325-347
18. Pavlov I., Excerpts from The Work of the Digestive Glands (Thompson W. translation), American Psychologist, V. 52 (9), 1997, P.936–940
19. Tesauro G., TD-Gammon - a Self-Teaching Backgammon Program, Achieves Master-Level Play, Neural Computation Vol. 6 (2), 1994, P.215–219
20. Cover T., Universal portfolios, Mathematical Finance Vol. 1 (1), 1991, P.1-29
21. Barto A. and Anandan P., Pattern-recognizing Stochastic Learning Automata, Vol. SMC-15 Issue 3, 1985, P.360-375
22. De Bondt W. and Thaler R., Does the Stock Market Overreact?, Journal of Finance, Vol. 40, 3, 1985, P.793–805
23. Sutton, R. and Barto A., Toward a Modern Theory of Adaptive Networks: Expectation and Prediction, Psychological Review Vol. 88 (2), 1985, P.135–170
24. Fukushima K., Cognitron: A self-organizing Multilayered Neural Network, Biological Cybernetics Biological Cybernetics V. 20, 1975, P. 121–136
25. Klopf H., A Comparison of Natural and Artificial Intelligence, ACM SIGART Bulletin Issue 52, 1975, P.11–13
26. Bellman R., A Markovian Decision Process, Journal of Mathematics and Mechanics, V. 6, No. 5, 1957, P.679-684
27. Walter W., An Imitation of Life, Scientific American Vol. 182, No. 5, 1950, P.42-45
28. Hull C., Simple Trial and Error Learning: A Study in Psychological Theory, Psychological Review, V. 37 (3), 1930, P.241–256

Unpublished Research

1. Guo Y., Fu. X., Shi Y., Liu M., Robust Log-Optimal Strategy with Reinforcement Learning, arXiv:1805.00205, 2018
2. Jiang Z., Xu D., Liang J., A Deep Reinforcement Learning Framework for the Financial Portfolio Management Problem, arXiv:1706.10059, 2017
3. Dochow R., Online Algorithms for the Portfolio Selection Problem, PhD Thesis, Saarland University, Germany, 2015
4. Watkins C., Learning from Delayed Rewards, King's College, Cambridge, 1989
5. Widrow B., Hoff M., Adaptive Switching Circuits, Stanford University, United States, 1960

3- المواقع الإلكترونية

1. بورصة عمّان: https://www.ase.com.jo
2. سوق دمشق للأوراق المالية: http://www.dse.sy