

مشروع PR2

التنبؤ بأسعار الأسهم باستخدام تقنيات تعلم الآلة

Machine Learning – Based Prediction

إعداد

حسين الصالح - محمد ديب

إشراف

دياسر خضرا

مقدم في 2020/10/05

الملخص

يقدم المشروع نظاماً للتنبؤ بأسعار الأسهم المستقبلية باستخدام تقنيات تعلم الآلة وذلك بالاستفادة من النماذج الاحصائية والخوارزميات والبرمجيات المتوافرة بالإضافة إلى عرض نتائج التنبؤ بشكل رسومي وتحديد دقة التنبؤ وفق معايير محددة، وقد عرض المشروع ضمن مقدمة وسبعة فصول مختلفة حيث تم في **المقدمة** استعراض المنطلق الذي بدأت منه أفكار التنبؤ بالمستقبل أما الفصول الرئيسية فتضمنت المراحل الأساسية للمشروع كالتالي: في **الفصل الأول** تم التعريف بالمشروع من حيث الهدف وشروط ومتطلبات النظام، ثم استعرضنا في **الفصل الثاني** الدراسة المرجعية التي تناولت منصة تداول مشهورة وأساليب التنبؤ التقليدية لأسعار الأسهم وعيوبها المحتملة، أما في **الفصل الثالث** جرى استعراض الدراسة التحليلية للنظام تحت عنوانين رئيسيين هما التخطيط وتحليل المتطلبات، حيث في التخطيط تم استعراض مبررات المشروع وتحليل الجدوى، والتخطيط لإدارة المشروع من حيث مراحله وموارده مروراً بإدارة المخاطر إلى تقسيم المهام بين فريق العمل ووضع خطة تنفيذ للمشروع (مخطط غانت) انتهاءً بمصفوفة المسؤوليات، وتحت عنوان تحليل المتطلبات تم استعراض المتطلبات الوظيفية وغير الوظيفية ومستخدمي النظام بالإضافة إلى حالات الاستخدام مفصلة مع جداولها، وفي **الفصل الرابع** كانت مرحلة التصميم، وجرى فيها استعراض تقنيات تعلم الآلة ومفاهيم مرتبطة بها بالإضافة إلى استكشاف البيانات ودراسة الخوارزميات والتقنيات، يلي ذلك **الفصل الخامس** الذي تم تخصيصه لشرح التنفيذ البرمجي وشرح الاجرائيات التي تم استخدامها واستعراضها لتحقيق وظيفة النظام، وفي **الفصل السادس** تم شرح مرحلتي التنفيذ والاختبار وفق سلسلة الاجرائيات، ومرحلة تنفيذ التقانات المستخدمة في المشروع بالإضافة إلى الاختبارات التي تشمل وظائف النظام لحساب الدقة والجودة والتحسينات الممكنة، وأما في الفصل الأخير وهو **الفصل السابع** فقد خصص لدليل الاستخدام حيث تم توضيح كيفية استخدام النظام لتحقيق المنفعة المرجوة منه واستعراض الواجهات الرئيسية، وفي **الخاتمة** تم اختتام التقرير بخاتمة عن العمل بمجمله ثم استعراض الآفاق المستقبلية والخيارات الممكنة لتطوير المشروع مستقبلاً.

الفهرس

3	الفهرس
6	المقدمة
7	تعريف المشروع
8	1- فكرة عامة عن المشروع
8	2- الهدف من المشروع
9	3- دفتر الشروط والمتطلبات
10	4- مراحل العمل في المشروع
10	1-4- توصيف المشكلة
10	2-4- معايير التقييم
11	الدراسة المرجعية
12	1- تقديم
12	2- تداول الأسهم والأوراق المالية
12	1-2- أساليب التحليل والتنبؤ التقليدية
13	2-2- المخطط البياني الخطي Line Chart
13	2-3- مخطط الأعمدة Bar Chart
14	4-2- الشموع اليابانية
17	3- منصة MetaTrader لتداول العملات
19	4- خوارزميات تنبؤ مستخدمة في الأبحاث
20	الدراسة التحليلية
21	1- التخطيط
21	1-1- إقلاع المشروع
22	2-1- إدارة المشروع
27	2- تحليل المتطلبات
27	1-2- مصدر المتطلبات
28	2-2- آلية جمع المتطلبات
28	3-2- المتطلبات الوظيفية
29	4-2- المتطلبات غير الوظيفية
30	5-2- مخطط حالات الاستخدام Use Case

التصميم	38
1- تقنيات تعلّم الآلة	39
1-1- التعلم الخاضع للإشراف Supervised Learning	39
2-1- التعلم غير الخاضع للإشراف Ununservised Learning	40
3-1- التعلّم المعزز Reinforcement learning	40
2- مفاهيم مرتبطة بتقنيات تعلّم الآلة	40
• البيانات	40
• المعلومات	41
• المعرفة	41
3- استكشاف البيانات	41
3-1- جدول قواميس البيانات	43
3-2- استكشاف البيانات رسومياً	43
4- الخوارزميات والتقنيات	44
5- النموذج المعياري وفق التراجع الخطي Linear Regression Model	45
6- المنهجية	46
6-1- المعالجة الأولية للبيانات	46
6-2- تدريب البيانات للنموذج المعياري باستخدام التراجع الخطي	48
6-3- تدريب البيانات لتطبيق الخوارزمية النهائية LSTM	48
7- التنفيذ	48
التحقيق البرمجي	50
1- إجراءات المعالجة	51
2- تسلسل الإجراءات	53
2-1- الخطوة الأولى	53
2-2- الخطوة الثانية	54
2-3- الخطوة الثالثة	55
2-4- الخطوة الرابعة	56
2-5- الخطوة الخامسة	57
2-5-5-2- توابع لظهور التنبؤ بسعر الاغلاق	59
التنفيذ والاختبارات	61
1 - التنفيذ	62

62	1-1- التقانات المستخدمة
62	2- إنجاز الإجراءات
63	1-2- تحصيل البيانات
63	2-2- المعالجة الأولية
66	2-3- بناء النموذج المعياري باستخدام التراجع الخطي
68	2-4- بناء نموذج التنبؤ باستخدام خوارزمية LSTM
70	2-5- بناء النموذج المحسن من خوارزمية LSTM
72	3- التحسين
72	4- النتائج
73	4-1- تطوير النموذج والتحقق من صحته
73	4-2- اختبار المتانة Robustness Check
73	4-3- التبرير
75	دليل الاستخدام
76	1- استعراض الواجهات الرئيسية
82	الخاتمة والآفاق المستقبلية
83	1- الخاتمة
83	2- الآفاق المستقبلية
84	المراجع
85	المراجع المستخدمة

المقدمة

يعد اكتشاف المستقبل وإدراكه من المزايا الخاصة بالوجود الإنساني، ومن الدلائل على تفرد الوعي البشري بخصائص استثنائية، ومن أجل هذا المستقبل المجهول قدّم أسلافنا الأضحيان وتفرّسوا السماء بحثاً عن الغيوم الماطرة أو النجوم التي تخيلوا أنها تحمل أقدارهم المستقبلية، زرعوا المحاصيل وخنوها خوفاً من المخاطر المحتملة عندما تعلّموا أن المستقبل يمكن التفاوض معه والاحتراز منه، لم يخرجوا بسفنهم في عرض البحار إذا ما قال لهم كاهن بحار أن البحر موعود بعواصف جيّارة، وكان التنبؤ بالمستقبل هو الموهبة الفريدة والمهارة التي رغبنا بامتلاكها بشدة حتى نفرض سيطرتنا على كل ما هو خارج نطاق معرفتنا.

وفي عالم اليوم الذي يرنو للعقلانية واتباع الطرق المنطقية والعلمية ما يزال المستقبل يثير اهتمامنا إما لمنافع معينة أو لاستباق أخطار متوقعة أو لإشباع غريزة الفضول نحو المجهول وتعلّم المزيد عن وجودنا وتجربتنا، ولكننا رغم استمرار اعتمادنا على الماضي واستخدام البيانات والمعلومات التي ورثناها للتنبؤ بالمستقبل، إلا أننا قد تجاوزنا مرحلة التكهنات العشوائية والافتراضات الميتافيزيقية التي لطالما كانت المحرك الأساسي والدافع الأولي في عملية التنبؤ عند أسلافنا، فاليوم لا يجب أن تقدّم قرباناً عند معبد هرمي الشكل للتنبؤ بالمستقبل، بل يكفي أن تزور مكتباً مكعب الشكل يجلس في وسطه باحث في مجال الذكاء الصناعي خلف شاشة وتطلب منه أن يزودك بالحكمة المفيدة للتعامل مع قضايا مستقبلية ما.

حكمة التنبؤ بالمستقبل اليوم أصبحت خاضعة للمنطق الاستقرائي وتعتمد على استخدام الأدوات الرياضية المتاحة وبشكل رئيسي الإحصاء والاحتمالات، والأهم من كل هذا هو البيانات التاريخية الضخمة التي أصبحت متوفرة بشكل يسهل الوصول إليه ومنظمة بطريقة تسهل التعامل معها، وتعد تقنيات الذكاء الصناعي مثل تعلّم الآلة والتعلّم العميق من الأدوات الجيدة التي يتم استخدامها في العديد من التطبيقات المتعلقة بالتنبؤ بالأحداث المستقبلية في المجال الاقتصادي والمالي والبنكي، وفي القضايا الاجتماعية مثل تحليل بيانات مرتبطة بمجتمع ما والحصول على معلومات مفيدة للتعامل مع مشكلة أخرى، وفي مجالات التسويق والاستثمار وفي المجال الصحي كما برز استخدامها مؤخراً بشكل مفيد منذ بداية الوباء العالمي كوفيد-19 وغيرها من التطبيقات.

تعتبر تقنيات تعلّم الآلة من أهم التقنيات المستخدمة حالياً في مجال الذكاء الصناعي وتطبيقاته المتعددة ولعل التطور التقني الحديث وتطوير مكتبات برمجية سهلت عمليات استثمار الخوارزميات المتعلقة بذلك ساهمت بانتشار هذه التقنيات و التوسع الكبير في تطبيقاتها.

ومن هذا المنطلق جاءت فكرة الاستفادة من البيانات الكبيرة واستخدامها في تطبيق خوارزميات تعلّم الآلة في عملية التنبؤ بسلوكيات محددة لأسعار الأسهم في أسواق التداول وتقييم النتائج، والوصول إلى تحقيق نتيجة تنبؤ أفضل مع كل مرحلة وباستخدام نماذج محسّنة، بالإضافة إلى التمثيل الرسومي، في خطوة أولى لبناء برنامج دعم للمستثمرين والمحليلين الاقتصاديين والشركات المدرجة في هذه الأسواق.

الفصل الأول

تعريف المشروع

1- فكرة عامة عن المشروع

البورصة أو سوق الأوراق المالية هي سوق لكنها تختلف عن غيرها من الأسواق، فهي لا تعرض ولا تملك في معظم الأحوال البضائع والسلع، فالبضاعة أو السلعة التي يتم تداولها بها ليست أصول حقيقية بل أوراق مالية أو أصول مالية، وغالباً ما تكون هذه البضائع أسهم وسندات، وتعتبر أسواق تداول الأسهم (البورصة) من الأعمدة الأساسية في نظم السوق المفتوح الرأسمالية، حيث توفر هذه الأسواق وسيلة لتقييم أسعار أسهم شركة ما، وطريقة لشراء وبيع الأسهم وتلعب دوراً أساسياً في تقييم صحة اقتصاد ما أو أداء شركة ما.

والبورصة سوق لها قواعد قانونية وفنية تحكم أدائها وتحكم كيفية اختيار ورقة مالية معينة وتوقيت التصرف فيها وقد يتعرض المستثمر غير الرشيد أو غير المؤهل لخسارة كبرى في حال قيامه بشراء أو بيع الأوراق المالية في البورصة لأنه استند في استنتاجاته في البيع أو الشراء على بيانات خاطئة أو غير دقيقة أو أنه أساء تقدير تلك البيانات، ولطالما سعت الشركات الاستثمارية وصناديق التحوط (Hedge funds) وحتى الأفراد، إلى استخدام النماذج المالية في سبيل الوصول إلى فهم أعمق لسلوك السوق وتحقيق أرباح استثمارية أو إنجاز صفقات رابحة.

ولكن الثروة الحقيقية هي المعلومات التي تكمن على شكل بيانات متوافرة وتمثل أسعار الأسهم التاريخية وبيانات أداء الشركات، والتي تعتبر بيانات ملائمة ليتم معالجتها من قبل خوارزميات تعلم الآلة (Machine learning).

2- الهدف من المشروع

السؤال الذي نطرحه من خلال هذا المشروع هو: هل يمكن التنبؤ بأسعار الأسهم في أسواق التداول باستخدام تقنيات تعلم الآلة؟

عادة ما يقوم المستثمرون باتخاذ قراراتهم بناء على معرفة متراكمة من خلال تحليل البيانات وقراءة الأخبار ودراسة تاريخ الشركة والمجال الصناعي الذي تعمل به والاطلاع على العديد من البيانات الأخرى المختلفة وكل هذا جزء من عملية التنبؤ التي يقوم بها المستثمر قبل أن يتخذ قراره في الاستثمار.

أما النظريات السائدة تقول أن أسعار الأسهم عشوائية تماماً ولا يمكن التنبؤ بها ولكن هذا يطرح تساؤلات عدة أهمها لم قد تقوم الشركات الاستثمارية الكبرى مثل (مورغان ستانلي - Morgan Stanley) و(سيتي غروب - Citigroup) بتوظيف محللين بيانات كمية بهدف بناء نماذج تنبؤ تساعد صناعات القرار.

والصورة النمطية عن أسواق التداول أنها قاعة ضخمة مليئة بالشاشات التي تعرض أرقاماً متتالية يركض فيها رجال غاضبون ومشتغلين بالأدريين ويصرخون عبر الهاتف ويحلّون ربطات عنقهم، هذه الصورة السينمائية لم يعد لها وجود في عالم اليوم، وغالباً ما يجلس هؤلاء الرجال بهدوء أمام خبراء تقنيات تعلم الآلة ويراقبون البيانات عبر شاشات حواسيبهم، وفي الواقع يتم حالياً تنفيذ حوالي 70% من أوامر الشراء والبيع في سوق تداول (وول ستريت - Wall Street) من قبل البرمجيات، هذا العصر الذي نعيش فيه هو عصر الخوارزميات.

يهدف هذا المشروع إلى استخدام نماذج تقنيات تعلم الآلة وخوارزمية الشبكات العصبونية (الذاكرة طويلة وقصيرة المدى LSTM - Long-Short Term Memory) بهدف التنبؤ بأسعار الأسهم في سوق التداول. إن استخدام الشبكات العصبونية المتكررة (Recurrent neural networks (RNNs) مفيد عند التعامل مع بيانات مدمجة في إطار زمني أي سلاسل زمنية (مثل أسعار الأسهم) ولكن أحدث الأبحاث تظهر أن استخدام خوارزمية شبكات LSTM هي الأكثر شيوعاً وفائدة من أنواع الـ RNNs. سوف يتم استخدام مكتبة Keras لبناء نموذج LSTM للتنبؤ بأسعار الأسهم بالاعتماد على البيانات التاريخية لسعر الإغلاق وحجم التداول وتمثيل البيانات رسوماً لكل من الأسعار التي تم التنبؤ بها عبر الوقت والمعايير الأفضل للنموذج.

3- دفتر الشروط والمتطلبات

يشترط في النظام المراد بناؤه أن يحقق هدف المشروع من خلال ما يلي:

- 1- إمكانية تحصيل مجموعات البيانات الكبيرة
- 2- تخزين البيانات بحيث يمكن إعادة استخدامها
- 3- تطبيق معالجة أولية على البيانات
- 4- إمكانية تمثيل البيانات رسوماً
- 5- تقسيم البيانات المتوفرة لمجموعتين للتدريب والاختبار
- 6- تطبيق نموذج تنبؤ
- 7- إمكانية عرض النتائج لنموذج التنبؤ
- 8- تمثيل نتائج التنبؤ رسوماً مقارنة بالنتائج الفعلية
- 9- حساب معدل الخطأ ودقة النموذج

أما بالنسبة لمستخدم النظام:

- 1- إمكانية إجراء عملية التنبؤ
- 2- عرض نتائج عملية التنبؤ ودقة النموذج
- 3- عرض التمثيل الرسومي للبيانات

4- مراحل العمل في المشروع

4-1- توصيف المشكلة

المشكلة التي يهدف هذا المشروع إلى حلها هي التنبؤ بشكل دقيق هي التنبؤ بأسعار الإغلاق المستقبلية لأسهم شركة ما ضمن فترة زمنية معينة في المستقبل.

سوف نستخدم خوارزمية الشبكات العصبونية الذاكرة طويلة وقصيرة الأمد - Long-Short Term Memory LSTM وتسمى اختصاراً للتنبؤ بأسعار إغلاق أسهم S&P500 شركة (ألفابت - Alphabet) باستخدام مجموعة بيانات dataset من الأسعار التاريخية السابقة.

الأهداف:

- استكشاف أسعار الأسهم
- تطبيق نموذج معياري باستخدام التراجع الخطي Linear regression
- تطبيق خوارزمية LSTM باستخدام مكتبة Keras
- تطبيق نموذج محسّن من خوارزمية LSTM
- مقارنة النتائج وتثبيت التقرير

4-2- معايير التقييم

من أجل قياس دقة هذا المشروع سوف أستخدم متوسط مربع الخطأ MSE وجذر متوسط مربع الخطأ RMSE لحساب الفروقات بين القيم التي تم التنبؤ بها والقيم الحقيقية للأسهم المستهدفة وفق سعر الإغلاق وقيمة (الدلتا) لحساب الفرق بين النموذج المبدئي (التراجع الخطي) والنموذج الرئيسي (خوارزمية التعلم الآلي)

الفصل الثاني

الدراسة المرجعية

1- تقديم

سنتناول في هذا الفصل معلومات أوسع عن طبيعة التداول في أسواق البورصة وكيف يقوم الخبراء والمستثمرون باتخاذ قراراتهم بطبيعة الحال، مما يمهّد لنا الطريق في إنجاز نظام التنبؤ كي يلبي احتياجات هذا المشروع، بالإضافة إلى دراسة منصات التداول الشهيرة حالياً للاستفادة منها في إثراء المحتوى ودراسة الخوارزميات المختلفة المستخدمة للتنبؤ بأسعار الأسهم في سوق التداول.

2- تداول الأسهم والأوراق المالية

تداول الاسهم هو عملية بيع و شراء الأسهم في شركة مدرجة في البورصة من خلال وسيط للتداول، على أمل تحقيق ربح من تغير الأسعار صعوداً أو نزولاً وهو ما يعرف بالمضاربة، أو بغرض الاستثمار طويل المدى في شركة تبدو واعدة وتكسب ثقة المستثمرين، عادة ما يعتبر انتقال الشركة من كونها شركة مساهمة محدودة إلى إعلان الاكتتاب العام وإدراج الأسهم في البورصة أمام جميع المستثمرين للتداول، نقلة نوعية للشركة. يتم تحديد سعر سهم معين من خلال العدد الإجمالي للأسهم التي أنشأتها الشركة، وعادةً ما يتم قياسها بعملة سوق الأسهم المدرجة فيها، على سبيل المثال، جنيه إسترليني في المملكة المتحدة ويورو في أوروبا والين في اليابان والدولار الأمريكي في الولايات المتحدة.

تماشياً مع قانون العرض والطلب، عندما يكون عدد المتداولين الذين يرغبون في شراء أسهم شركة أكبر من عدد أولئك الذين يريدون بيعها، يرتفع سعر سهمها عادة، وعلى العكس من ذلك عندما يكون هناك عدد المتداولين الذين يرغبون في بيع أسهم شركة أكثر من المتداولين الذين يرغبون في شرائها، يميل سعر السهم إلى الانخفاض.

ويشار إلى مستوى سوق الأسهم بما يسمى نقطة، ويتم إحصاء النقاط للخسارة والارتفاع بما يسمى سعر الإغلاق للسوق في اليوم.

2-1- أساليب التحليل والتنبؤ التقليدية

يعتمد المتداولون (المتعاملون) عموماً على أسلوبين في اختيار الأسهم، التحليل الفني وهو أسلوب يمكن من فحص الأوراق المالية وفقاً لتطور سعرها وتحركات السعر التاريخية وباستخدام الرسوم البيانية وذلك لتحديد توقيت التصرف بمعنى متى يتم شراء السهم أو بيعه أو الاحتفاظ به والتحليل الأساسي وهو فحص للقوائم المالية وذلك وصولاً إلى القيمة الحقيقية للسهم بحيث يساعد التحليل الأساسي في التعرف على الأسهم ذات الخلل السعري أي المُسَعَّرَة بأقل أو بأعلى من قيمتها الحقيقية، ولا يمكن القول أن التحليل الفني أفضل من التحليل الأساسي أو العكس ولكن المستثمر يحتاج للتحليل الأساسي لاختيار الأوراق المالية الجيدة ويحتاج للتحليل الفني للمساعدة في تحديد توقيت اتخاذ القرار، وجدير بالذكر أن أنصار التحليل الفني يرون أنهم هم الأدق والأفضل حيث يتنبؤون بالمستقبل باستخدام فكرة التاريخ يعيد نفسه بينما يعتبروا التحليل الأساسي تحليل قاصر لكونه

يعتمد على أحداث ماضية تاريخية، ويتغير سعر السهم في سوق الأسهم كنتيجة مباشرة لتغير نسب العرض والطلب على هذا السهم أو ذاك، ففي حالة الإقبال الشديد على الشراء فإن طلبات البيع رخيصة الثمن سوف تنفذ، وتبدأ الطلبات الأكثر سعرا بالظهور ويبدأ معها السعر بالارتفاع تدريجيا، وهذا على عكس مايجري في حال الإقبال على البيع.

هذا لا ينفي أهمية الأوضاع السياسية والاقتصادية والعلاقات بين الدول بالإضافة إلى الأخبار اليومية للأشخاص والأفراد المؤثرين من أصحاب الشركات أو قادة الأحزاب السياسية أصحاب التوجهات التي يتوقع بأنها تطرح أفكاراً راديكالية، فإذا تمكن حزب سياسي ينتهج سياسيات ترغب بفرض ضرائب قاسية على الشركات من تحقيق الفوز مثلاً، سوف يشكل هذا ضربة لسوق الأوراق المالية والشركات في هذه الدولة والدول المتعاملة معها، أو إذا قام المدير التنفيذي لشركة ما بفعل يثير ضجة الإعلام فهذا حتماً سوف ينعكس على أسهم الشركة.

2-2- المخطط البياني الخطي Line Chart

المخطط الخطي هو أحد أنواع المخططات البيانية التي تعرض المعلومات كسلسلة من نقاط البيانات المتصلة بواسطة خطوط مستقيمة، وهو النوع الأساسي من أنواع المخططات البيانية شائعة الاستخدام في كافة المجالات وفي مجال تحليل أسواق البورصة بالطبع، ويتم إنشاء هذا المخطط بوصل سلسلة من النقاط التي تمثل مقاييس مفردة بواسطة قطاعات خطية، فقد تكون النقاط تمثل إما سعر الإغلاق أو سعر فتح التداول أو أعلى قيمة تحققت في اليوم أو أقل قيمة، ومثالاً عليه الشكل 1.



الشكل (1) المخطط البياني

2-3- مخطط الأعمدة Bar Chart

ويسمى أيضاً مخطط بياني شريطي هو رسم بياني تمثل فيه البيانات بأشرطة مستطيلة طول كل منها يتناسب مع القيمة التي يمثلها، يستخدم شريط الرسوم البيانية للمقارنة بين قيمتين أو أكثر من القيم التي أخذت حسب زمن أو ظروف مختلفة، فيمكن استخدامه لتمثيل (الافتتاح، الإغلاق، الحد الأدنى، الحد الأعلى) لأسعار الأسهم حسب

الزمن، ويمكن أن يكون الشريط في المجموعات البيانية الصغيرة خطأ أفقياً أو يمكن أن يكون شاملاً لنقطة العرض ومثالاً عليه في الشكل (2)



الشكل (2) مخطط الأعمدة

4-2- الشموع اليابانية

هي من أشهر طرق التحليل الفني للأسهم أو بورصة العملات، يرجع تاريخ الشموع اليابانية إلى حدود عام 1600م حيث كان اليابانيون يتاجرون في سوق الأرز، وقام مضارب يسمى هوما مونهيسا بابتكار تقنية فنية لتمثيل سعر الأرز وتحليله وسميت هذه التقنية بالشموع اليابانية، تعد نماذج الشموع اليابانية اليوم أداة لا تقدر بثمن من مجموعة أدوات المتداولين الحديثة، حيث يتم استخدامها لوصف تحركات الأسعار لأسهم أو عملة أو أداة مشتقة سائلة معينة مثل العقود الآجلة، وتعتبر الشموع اليابانية تقنية فريدة من نوعها لقراءة رد فعل السعر والتعرف على نفسية المستثمرين من خلال مجموعة من النماذج التي تعطي إشارات مبكرة ذات قيمة كبيرة للمحللين.

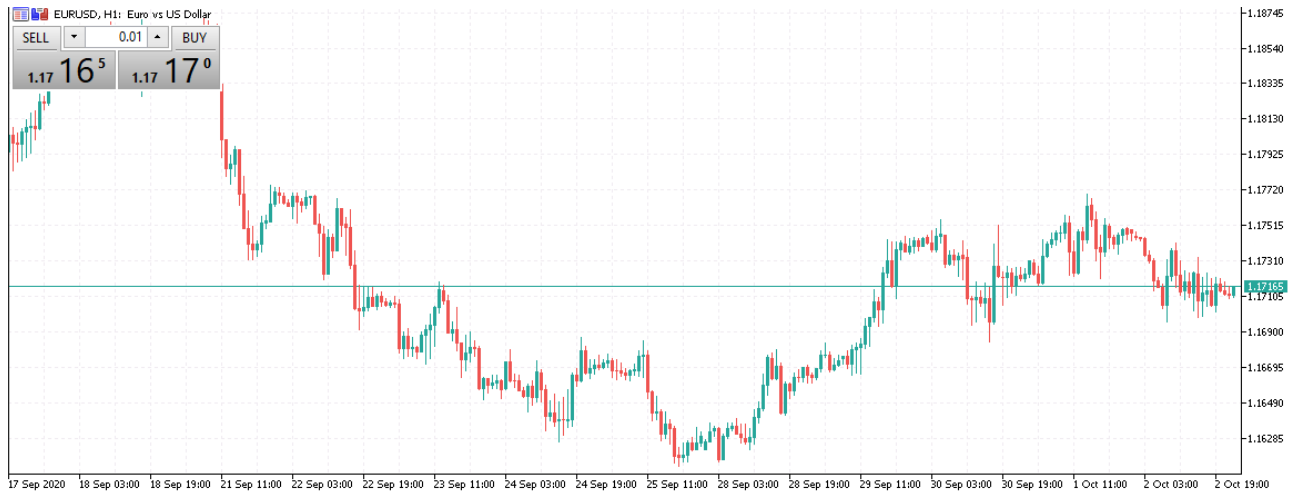
ومقارنة الشموع اليابانية برسم الأعمدة ظالمة جداً، فالشموع اليابانية تتفوق إلى حد كبير فهي توضح نفس البيانات التي توضحها الأعمدة (الافتتاح، الإغلاق، الحد الأدنى، الحد الأعلى) ولكن في علاقة بين هذه الأسعار لتشكل نماذج تعطي انطباع صادق عن نفسية المستثمرين ومثالاً على الشموع اليابانية في الشكل (3).



الشكل (3) الشموع اليابانية

كقاعدة عامة، إذا كانت الأسعار تتجه للإرتفاع خلال تكوين الشمعة، يتحول جسم الشمعة إلى اللون الأبيض أو يترك فارغاً، وإذا انخفضت الأسعار يظل الجسم باللون الأسود أو أي لون آخر وإن كانت الشموع المتزايدة في هذا الرسم البياني لها لون الخلفية، وبهذا يسمح تحليل الشمعة اليابانية للتاجر أن يفهم كيف تغيرت الأسعار خلال فترة زمنية معينة.

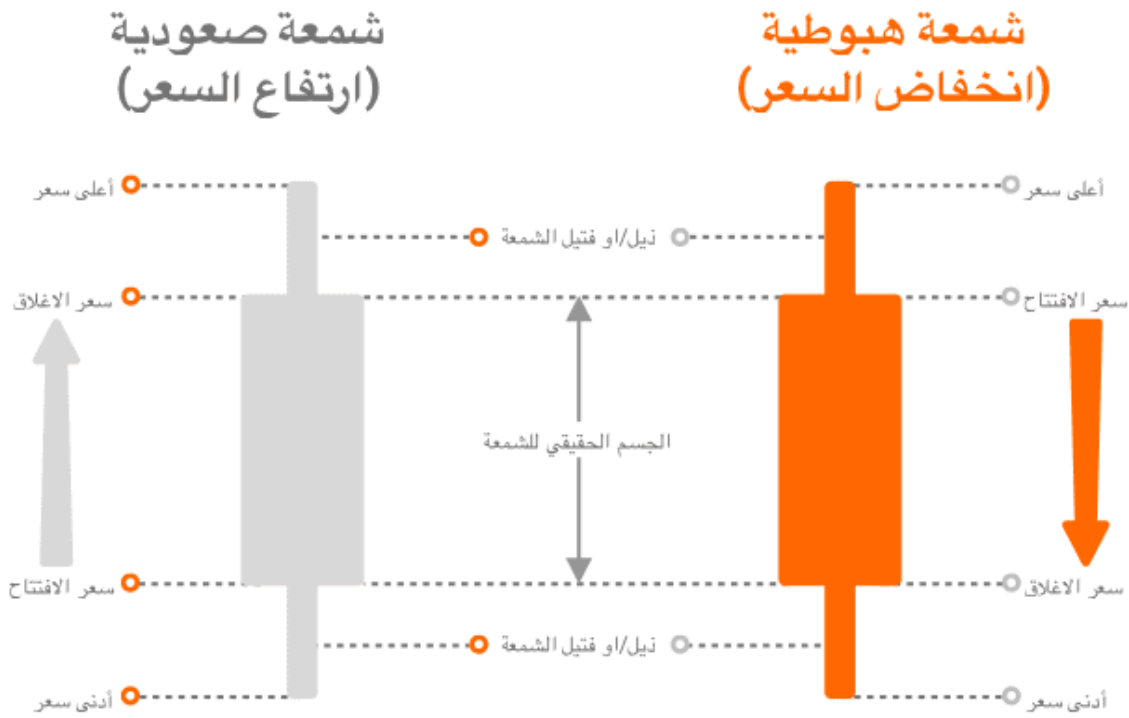
معايير الانخفاضات والارتفاعات تمثل مجموعة كاملة من تقلبات الشمعة وتشير إلى تغيرات حقيقة في العملات أو الأسهم أو أسعار أصول أخرى، ويكتسب هذا الأمر أهمية خاصة في الرسوم البيانية على المدى طويل الأجل، حيث التغيرات في تداولات اليوم يتم عرضها جميعاً في وقت الإغلاق.



الشكل (4) مثال على الشموع اليابانية لزواج تداول اليورو والدولار

ولتحليل السوق بنمط الشموع اليابانية يجب الأخذ بعين الاعتبار كل شمعة حيث تعطي المتداول معلومات معينة، بالإضافة إلى الأشكال التي قد تشكلها الشموع على طول السلسلة الزمنية مما يوحي بوجود حركة ما في السوق. على سبيل المثال:

يشير الظل الطويل إلى حركة قوية في الاتجاه الذي تم توجيهه إليه، لكن بعد أن قامت مجموعة من المتداولين بإعادة توجيه الحركة إلى الجانب الآخر يمكن اعتبار هذا بمثابة إشارة لتغير الاتجاه، والصفقات المفتوحة في الاتجاه المعاكس حيث يشير الظل، ومع ذلك أكثر الإشارات فاعلية تعطى من قبل الشموع. وهذه بعض النماذج التي يمكن تشكيلها من قبل واحدة أو مجموعة كاملة من الشموع، وتنقسم هذه الأنماط إلى مجموعات يشير بعضها إلى إمكانية حدوث انعكاس في الاتجاه بينما تشير غيرها إلى استمراره. أنماط الشموع هي تشكيلات تداول فنية تساعد على تصور حركة سعر الأصول ذات السيولة (مثل الأسهم، العملات الأجنبية، العقود الآجلة، وغيرها). يتم استخدامها من قبل المتداولين لتحديد توقيت نقاط الدخول والخروج بشكل أفضل، وافترضياً تعرض مخططات الشموع اليابانية فواصل زمنية ليوم واحد، هذا يعني أنه في حال كنت تراقب مخططاً في إطار زمني لشهر واحد، فمن المحتمل أن ترى عدد 20 من الشموع اليابانية ومع ذلك، يمكن أن تمثل مخططات الشموع اليابانية فترات زمنية أطول أو أقصر من يوم واحد.



الشكل (5) تفصيل الشمعة الهبوطية والشمعة الصعودية

تتكون الشمعة الواحدة من أربعة مكونات مختلفة تمثل أجزاء مهمة من معلومات التسعير ليوم التداول وهي الافتتاح والإغلاق والارتفاع والانخفاض، أول سعرين هما الفتح والإغلاق ويتم تمثيلهما بالجسم السميك للشمعة ويشير فتيل الشمعة و ذيلها إلى نقاط السعر العليا والدنيا التي تم تسجيلها على مدار اليوم ونظراً لكونها مليئة بالمعلومات القيمة، فإن أنماط الشموع اليابانية أصبحت واحدة من أكثر أدوات التداول المفضلة للتنبؤ.

3- منصة MetaTrader لتداول العملات

رابط تحميل المنصة لكافة أنظمة الحاسوب والهواتف المحمولة: <https://www.metatrader4.com/>

تعد منصة MetaTrader أحد أشهر البرمجيات المستخدمة في سوق التداول، ومستخدمة بكثرة بين المضاربين أو المحللين الفنيين، بالإضافة إلى أنها توفر خيارات متعددة للتداول مثل الذهب والنفط وغيرها، ولكنها تعتبر منصة رئيسية للتداول بين العملات الأجنبية أو ما يعرف بسوق (فوركس) وهي مختلفة عن سوق الأوراق المالية.

1-1-3 الفرق بين سوق الأوراق المالية وسوق تداول العملات

وفيما يلي نعرض الاختلافات الرئيسية بين تداول الأسهم وتداول العملات (الفوركس):

- 1- في سوق البورصة يتم تداول الأوراق المالية وسندات وأسهم الشركات، أما في سوق الفوركس يتم تداول العملات، مثلاً الدولار الأمريكي مقابل اليورو، أو الريال السعودي مقابل الين الياباني، وغيرها من أزواج العملات التي يسعى المتداولون لتحقيق مكاسب من خلال البيع والشراء والتحويل بينها.
- 2- حجم التداول: لدى سوق الفوركس حجم تداول أكبر من سوق الأوراق المالية بما أن المتداول هو العملات النقدية ذاتها المستخدمة لإنجاز الصفقات في سوق الأوراق المالية بالإضافة إلى عمليات التحويل بين العملات التي تتم خارج سوق الأوراق المالية مثل الصفقات بين الدول.
- 3- تنوع الأدوات: هناك آلاف الأسهم للاختيار من بينها في مقابل بضع عشرات من أزواج العملات.
- 4- تأثير الأسعار: تتأثر أسعار الأسهم بشكل أساسي بالعوامل الداخلية مثل التقارير المالية وأحداث الشركات الأخرى (أرباح الأسهم وتجزئة الأسهم وما إلى ذلك)، في حين تتأثر أزواج العملات الأجنبية في الغالب بالعوامل الخارجية مثل التطورات السياسية والاقتصادية الإيجابية أو السلبية بين الدول.
- 5- تقلبات السوق: يمكن أن تتذبذب أسعار الأسهم بشكل كبير من يوم لآخر وتكون تقلباتها أكثر حدة بشكل عام من تلك الموجودة في أسواق (الفوركس) ولكنها تبقى الخيار الأمثل للاستثمار بعيد المدى، أما الاستثمار في (الفوركس) يمكن أن يكون مربحاً على المدى القريب.

2-1-3 مزايا المنصة:

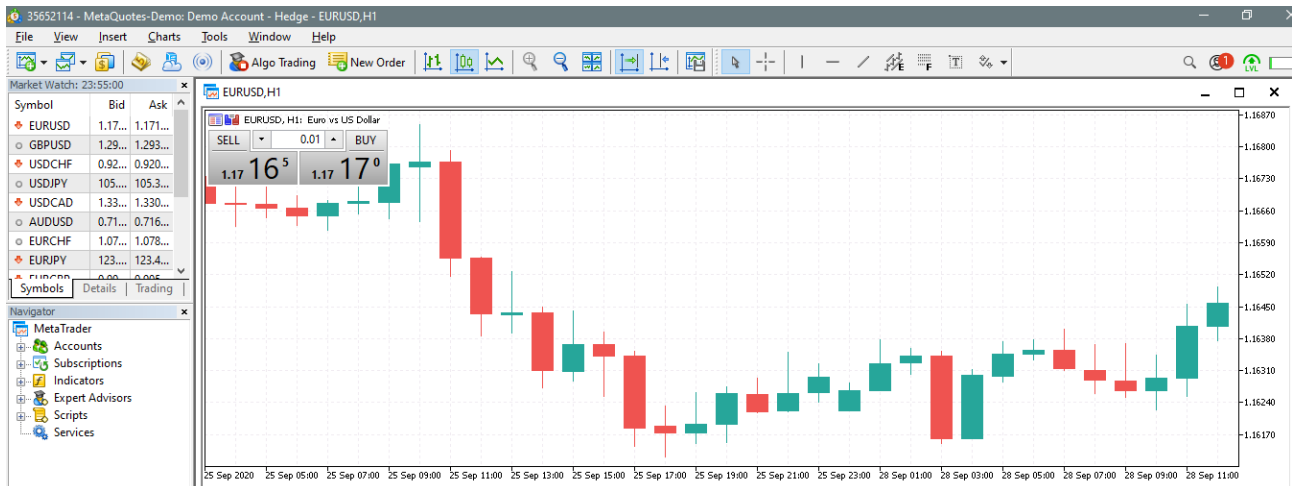
توفر منصة MetaTrader مزايا عديدة للمتداولين:

- خيارات متعددة للتداول ما بين عملات أو سلع
- توفر معلومات تفصيلية حيّة (كما يظهر في الشكل أدناه) عن الأسعار في كل لحظة
- تتيح للمتداول استعراض البيانات بشكل يساعده على اتخاذ قراراته
- تقدم خصائص متعلقة بأوامر الشراء والبيع مثل نفذ أمر البيع بشكل تلقائي إذا ارتفع السعر إلى أدنى من رقم معين أو عند تحقيق ربح معين

- عرض البيانات لأي زوج أو مجموعة من أزواج العملات التي يرغب بها بالإضافة إلى إمكانية اختيار طريقة عرض البيانات التي يفضلها المستخدم
- تطبيق بعض المؤشرات مثل الوسطي المتحرك أو استخدام أدوات لتحديد اتجاه السوق
- تدعم كافة أسواق تداول العملات من طوكيو شرقاً وحتى نيويورك غرباً وبالزمن الحقيقي



الشكل (6) واجهة منصة MetaTrader



الشكل (7) الشموع اليابانية في منصة MetaTrader

3-1-3 عيوب المنصة:

رغم أنها من أفضل المنصات للتداول إلا أنها تعاني من بعض العيوب:

- عدم إمكانية إجراء تنبؤ مستقبلي بأي شكل
- البيانات المعروضة تمثل فقط سوق تداول أسعار العملات ولا تشمل سوق الأوراق المالية
- يعتمد المستثمر على خبرته ومهارته في الاستثمار أو يقوم باستئجار خدمات تنبؤ مدفوعة

4- خوارزميات تنبؤ مستخدمة في الأبحاث

هناك العديد من الأعمال ذات الصلة التي يتم إجراؤها على سوق الأوراق المالية وما زالت مستمرة أيضاً وبعضها مذكور أدناه:

- 1- توقع أسعار الأسهم باستخدام نموذج ARIMA
نموذج متوسط متحرك متكامل ذاتي الانحدار وهذا النموذج يستخدم للتنبؤات قصيرة المدى.
- 2- توقع أسعار الأسهم باستخدام LSTM و RNN و CNN نموذج النافذة المنزلق
يعمل هذا النموذج باستخدام مفاهيم التعلم العميق ويستخدم كل من الخوارزميات الخطية وغير الخطية.
- 3- التنبؤ بسعر السهم باستخدام الشبكة العصبية مع السوق للمؤشرات المهيمنة
يعتمد هذا النموذج على ANN (شبكة عصبية اصطناعية) والتي تستخدم كأداة للتعيين والتنبؤ المالي.
- 4- دمج metaheuristics والشبكات العصبية الاصطناعية لتحسين توقع سعر السهم
يستخدم هذا النموذج تكامل metaheuristics و ANN لتحسين التنبؤ بالأسعار وكل من طوبولوجيا ANN وعدد المدخلات يكون محسناً
- 5- التنبؤ بأسعار الأسهم باستخدام تعزيز التعلم
يقترح هذا النموذج طريقة تستخدم التعلم المعزز للتنبؤ بأسعار سوق الأوراق المالية على المدى الطويل.

الفصل الثالث

الدراسة التحليلية

1- التخطيط

يعتبر التخطيط من أهم الأنشطة الإدارية لأي مشروع حيث يتوجب على إدارة المشروع الاهتمام به اهتماماً خاصاً لأنه مفتاح النجاح للمشروع حيث يتم تحديد مراحل المشروع و التخطيط لها من حيث الزمن و الموارد المادية و البشرية بما يؤمن تحقيق أهداف المشروع بالوقت المحدد ويعتبر مرجعاً يمكن العودة إليه عند حدوث المشاكل ووسيلة لتفادي حدوث أخطاء قدر المستطاع.

و حيث أن هدف المشروع هو تصميم نظام للتنبؤ بأسعار الأسهم باستخدام تقنيات تعلم الآلة يمكن ربطه مع نظام إدارة المؤسسات الاقتصادية والتجارية، وحسب الوقت والموارد والمخاطر المتوقعة والتي يمكن أن تحدث أثناء التنفيذ تم تحديد الخطوات الواجب اتباعها بوضوح للوصول للنتيجة المرجوة في نهاية المشروع.

1-1- إقلاع المشروع**1-1-1 الأسباب الموجبة للمشروع (مبررات المشروع)**

تتبع الحاجة الأساسية للتنبؤ بالتغيرات المستقبلية لأسعار الأسهم والأوراق المالية في سوق التداول من حاجة المستثمرين على مستوى الأفراد أو الشركات إلى التعامل بشكل صحيح مع تقلبات الأسعار التي لا بد من أن تحدث في هذا السوق، واستباق الفرص في تحقيق الأرباح أو على الأقل تجنب الخسائر إن أمكن، وفي الواقع فإن وجود كم هائل من المؤسسات البحثية وشركات تحليل البيانات الداعمة بالإضافة إلى الأقسام المتخصصة في الشركات والبنوك الاستثمارية التي تستقطب أفضل المحللين للعمل لديها يؤكد على أهمية هذا المجال.

بما أننا نعيش في عصر مليء بالتقنيات والإمكانيات والخوارزميات التي يتم تحسينها يومياً، فلا بد من الاستفادة من ذلك في سوق الأوراق المالية، فنحن نستخدم خوارزميات التعلم الآلي من أجل توقع اتجاه السوق وتقليل مخاطر رؤوس أموال المستثمرين في الاستثمار في سوق الأوراق المالية، والنمو الكبير الذي حققته التكنولوجيا والتقنيات الرقمية يزداد يوماً بعد يوم، ولهذا قررنا إنجاز هذا المشروع لوضع حجر إضافي في هرم المعرفة فوق ما سبقه من تجارب وأبحاث، وتقديم تنبؤات يُعتمد عليها ويمكن أن يستفاد منها المستثمر.

2-1-1 تحليل الجدوى

نبيّن فيما يلي بعض النقاط الأساسية التي تبين جدوى المشروع:

- جمع البيانات المتعلقة بأسهم شركة من أسواق التداول
- توفير وقت وجهد المحللين في إجراء العمليات الأولية للتنبؤ
- سهولة الوصول إلى التمثيل الرسومي للبيانات
- تأمين قدر معين من الموثوقية بالنتائج من خلال المعايير الإحصائية
- إثراء المجال المتعلق بالتنبؤ بأسعار الأسهم
- تقديم نماذج تنبؤ مختلفة ومقارنة النتائج

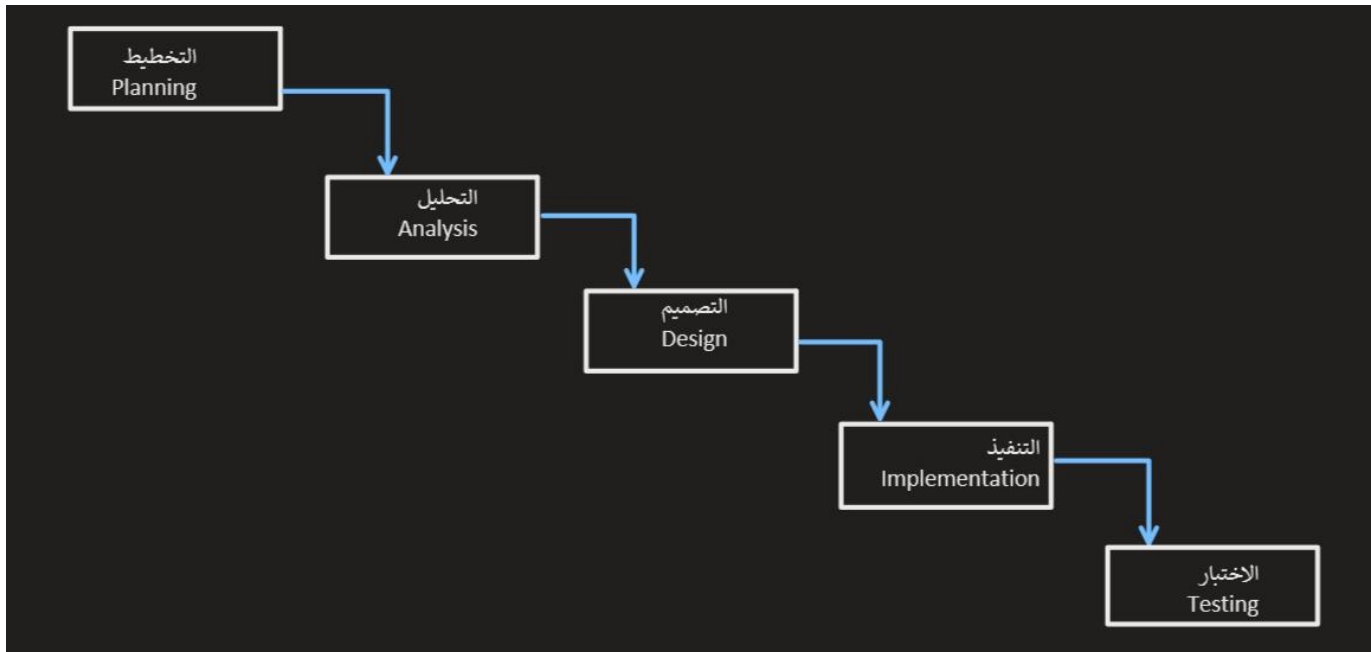
- تقليل عامل المخاطرة التي يتحملها المتداول أثناء عمليات البيع والشراء
- تقديم معلومات مفيدة مرتبطة بأداء أسهم الشركة
- تطبيق خوارزميات التعلم الآلي في بناء نماذج تنبؤ فعّالة

1-2- إدارة المشروع

تتمثل الإدارة الفعّالة للمشروع في امتلاك الرؤية العامة للمشروع بداية بالخطوة الأولى وحتى النتيجة النهائية، وتحديد الخطوات المطلوبة وتجميع التفاصيل في مهام واضحة ومحددة ومنفصلة ومتراصة، بحيث تستطيع الإدارة من بعدها أن تحدد إطاراً زمنياً لكل مهمة وأن تأخذ في عين الاعتبار الإطار الزمني المطلوب لإنجاز المشروع كاملاً بالإضافة إلى تمتعها بالمرونة للتغلب على المشاكل المتوقعة الحدوث أو غير المتوقعة، ووفق هذه الرؤية سوف يتم توزيع المراحل على الفترة الزمنية المتاحة وفقاً للمعلومات والمعطيات الحالية ومع اجراء مراجعات دورية للتأكد من إنجاز المهام في وقتها والمشروع النهائي بأنسب الطرق.

1-2-1 منهجية التطوير

وفقاً للرؤية الموضوعية نحو إنجاز المشروع يجب تحديد منهجية سير المهام والآلية التي يجب إتباعها في تحقيق الأهداف الجزئية وصولاً للهدف الكلي، ومن هذا المنطلق سوف يتم بناء هذا المشروع وتطوير نظام التنبؤ بالاعتماد على النموذج الشلالي WaterFall بما أن مراحل بناء النظام واضحة ومتسلسلة في معظمها ويمكن إتباعها تدريجياً كما هو موضح في الشكل (8).



الشكل (8) النموذج الشلالي Waterfall

2-2-1 موارد المشروع

- الموارد البشرية
 - فريق العمل: الطالب حسين الصالح، الطالب محمد ديب
 - المشرف على المشروع: الدكتور ياسر خضرا
- الموارد المادية
 - جهاز حاسوب عدد 2
- الموارد البرمجية
 - بيئة البرمجة Pycharm باستخدام لغة Python
 - برنامج Star UML لرسم المخططات
 - بيئة تطبيق المراحل Jupyter notebook
 - مكتبة تصميم الواجهات PyQt5, QT designer

3-2-1 إدارة المخاطر

يمكن تعريف الخطر في أي مشروع من المشاريع بأنه حدث أو ظرف غير مؤكد والذي ينتج عن حدوثه أثر سلبي مرتبط بشكل مباشر بالمشروع، والمخاطر التي تهدد المشروعات تعتبر ظاهرة طبيعية لكن الشيء غير الطبيعي هو عدم اهتمام الإدارة بحماية نفسها ومشروعها من المخاطر الكثير المحتملة وغير المتوقعة، ولا بد من إدارة المشروع أن تسعى إلى دراسة المخاطر والتهديدات الحقيقية (أو المحتملة) التي يمكن أن تواجهها حاضراً ومستقبلاً جاهدة إلى تخفيض مقدار الخسائر المحتملة إلى أدنى حد ممكن في حال استبعاد إمكانية إلغاء المخاطر بالكامل، بحيث تحافظ على أمان سير عملية مهمات تنفيذ المشروع وتراعي الجودة المطلوب تحقيقها. وبالتالي فإن إدارة المخاطر هي سلسلة من الإجراءات التي تهدف إلى الحد من وقوع الخطر وتقليل حجم الخسائر الناجمة عن وقوع الخطر وتوفير خطط طوارئ للتعامل مع المخاطر وإجراءات لتجاوزها بعد وقوعها.

وهذا يقتضي من إدارة المشروع:

- تقدير مدى الخطر المحتمل ونتائجه
- حجم الأثر السلبي لوقوع الخطر على المشروع
- اختيار أفضل الطرق والأساليب لتفادي وقوع الخطر
- استراتيجية تفادي الأثر السلبي بعد وقوع الخطر

الجدول (1) يوضح أنواع المخاطر المحتملة وحجم الأثر والطرق المتبعة لتفادي الضرر.

وصف الخطر	احتمال حدوثه	حجم الأثر على المشروع وتوصيفه	طريقة تفادي الأثر
1 تعطّل في الأجهزة أو حدوث خلل في تكامل إحدى البرمجيات أو المكتبات أو انقطاعات في شبكة الانترنت	متوسط	أثر كبير يسبب تأخير في تسليم المشروع وتخلّف عن إنجاز المهام في وقتها ورداءة في الأداء	صيانة الأجهزة دورياً والحرص على التنسيق في كل الخطوات المرتبطة بأدوات العمل وتأمين اشتراك انترنت احتياطي إضافي
2 عدم قدرة أحد أعضاء الفريق على إنجاز المهام الموكلة إليه لاسباب اضطرارية (مرضية أو حوادث)	ضئيل	أثر ضخم زيادة الخطر في عدم إتمام المهام في وقتها وتدني جودة العمل المنجز	إعداد خطة بديلة للتعامل مع حالات الطوارئ الاستعانة بخبرات خارجية إن أمكن
3 سوء التقدير للزمن اللازم لإنهاء المشروع	متوسط	أثر كبير يؤدي لانخفاض الجودة وتأخر في التسليم للمهام الجزئية والمشروع	التخطيط الجيد للزمن مراقبة وتتبع الخلل والمرونة في زيادة ساعات العمل
4 حدوث انقطاع مفاجئ وغير اعتيادي للتيار الكهربائي	متوسط	أثر ضخم يؤدي لتأخير إنجاز المهام وتعذر الالتزام بخطة العمل بشكل صحيح	توفير أجهزة محمولة بديلة تأمين مصادر بديلة للطاقة (بطاريات أو مولدات طاقة)
5 البعد الجغرافي لفريق العمل وعدم القدرة على التواصل المباشر	ضئيل	أثر ضئيل يؤدي لسوء تنسيق المهام بين فريق العمل وحدوث تضارب تنظيمي والتخلف عن الخطة الزمنية	الاعتماد على وسائل تواصل تسهل التنظيم كالماسنجر و Team Viewer
6 اختلاف الأفكار بين أعضاء الفريق والطرق المتبعة لتنفيذ المهام المتفق عليها	ضئيل	أثر كبير يؤدي لعرقلة تنفيذ خطة العمل والحاجة إلى إعادة توصيف المهام والمتطلبات مراراً وتكراراً	العمل على إيجاد نقطة التقاء وتقبل وجهة النظر المختلفة والاستعداد للحوار في سبيل إنجاح العمل النهائي للمشروع

الجدول (1) أنواع المخاطر المحتملة وطرق تفادي الأثر

سيتم تقسيم المشروع إلى ستة مراحل Phases هي كالتالي :

مرحلة التخطيط، مرحلة التحليل، مرحلة التصميم، مرحلة التنفيذ، مرحلة الاختبار، مرحلة الإعداد، مرحلة الإنهاء والتسليم، وتم توزيع المهام حسب الخطة الزمنية كما في الجدول (2)

Task Name	Duration	Start	Finish	Notes
الدراسة الأولية للمشروع	7 days	20/5/2020	27/5/2020	تحديد فكرة المشروع وهدفه وآليات إنجازه
التعرف على شروط المشروع	3 days	28/5/2020	31/5/2020	
مناقشة شروط المشروع مع المشرفين	2 days	1/6/2020	3/6/2020	
بدء المشروع	120 days			
مرحلة التخطيط للمشروع	6 days	4/6/2020	10/6/2020	
مرحلة التحليل	20 days			
تحديد وجمع المتطلبات	6 days	11/6/2020	17/6/2020	
التعرف على تقنية التعلم باستخدام الآلة	10 days	18/6/2020	28/6/2020	
رسم المخططات	7 days	29/6/2020	5/7/2020	
مرحلة التصميم	40 days			
سحب البيانات وتنظيمها	10 days	5/7/2020	15/7/2020	
linear نشاط تنفيذ regression	30 days	15/7/2020	15/8/2020	
مرحلة التنفيذ	45 days			
LSTM نشاط تنفيذ	15 days	16/8/2020	30/9/2020	
نشاط تصميم الواجهات	10 days	28/8/2020	7/9/2020	
نشاط اختبار الواجهات	3 days	8/9/2020	11/9/2020	

نشاط ربط وتكامل الكود	5 days	12/9/2020	17/9/2020	
نشاط كتابة تقرير المشروع	7 days	18/9/2020	25/9/2020	
نشاط تجهيز عرض تقديمي للمشروع PowerPoint	1 day	26/9/2020	27/2020	
نشاط تجهيز شرح فيديو للمشروع Demo	1 day	28/9/2020	29/9/2020	
نشاط تسليم المشروع	1 days	29/9/2020	30/9/2020	

الجدول (2) مخطط غانت لإنجاز العمل

5-2-1 مصفوفة المسؤوليات : RM

تتوزع المسؤوليات لإنجاز المشروع على الفريق العامل بأكمله و حسب طبيعة النشاطات المحددة بالفقرة السابقة وقد تمت المشاركة بين فريق العمل بجميع الخطوات لتنفيذ المشروع كما يلي:

- المسؤول (R) Responsible : وهو المسؤول عن تنفيذ النشاط وتحقيقه بشكل واقعي .

- المساعد (S) Support : وهو الذي يقوم بتقديم المساعدة والدعم للمسؤول

يوضح الجدول (3) مصفوفة المسؤوليات

الرمز	النشاط	الموقع	الوظيفي
1	الدراسة الأولية للمشروع	S	R
2	التعرف على شروط المشروع	R	S
3	مناقشة شروط المشروع مع المشرفين	R	S
4	بدء المشروع		
5	مرحلة التخطيط للمشروع	S	R
6	مرحلة التحليل		
7	تحديد وجمع المتطلبات	R	S

8	التعرف على تقنية التعلم باستخدام الالة	R	S
9	رسم المخططات	R	S
10	مرحلة التصميم		
11	سحب البيانات وتنظيمها	R	S
12	نشاط تنفيذ linear regression	R	S
13	مرحلة التنفيذ		
14	نشاط تنفيذ LSTM	R	S
15	نشاط تصميم الواجهات	S	R
16	نشاط اختبار الواجهات	S	R
17	نشاط ربط وتكامل الكود	S	R
18	نشاط كتابة تقرير المشروع	R	S
19	نشاط تجهيز عرض تقديمي للمشروع PowerPoint	R	S
20	نشاط تجهيز شرح فيديو للمشروع Demo	R	S
21	نشاط تسليم المشروع	R	S

الجدول (3) مصفوفة المسؤوليات

2- تحليل المتطلبات

تم في الفصل الأول توصيف متطلبات المشروع في فقرة دفتر الشروط والمتطلبات وفي هذا الفصل يتم تجميع المتطلبات الوظيفية وغير الوظيفية وتوصيفها بحيث تكون موجهة للمصممين والمطورين، كما يتم في هذه المرحلة تحديد الفاعلين في النظام وتوصيف حالات الاستخدام باستخدام مخططات UML

2-1- مصدر المتطلبات

ان جميع مشروعات تطوير النظام تحتاج إلى مشروعات موازية للتغيير ودائماً ما يكون لهذه المشروعات دورة حياة لها بداية ولكن ليس لها نهاية لأن التطوير والتغيير هو عملية مستمرة لا تتوقف.

ودائماً ما تكون البداية في هذه المراحل هي تحديد المتطلبات والتي غالباً ما يتم التعبير عنها في صورة عامة ومن منظور المستخدمين لنظام العمل، وخلال مراحل التطوير تصبح مهمة محلي النظم مع فريق التطوير تحويل هذه المتطلبات تدريجياً إلى مواصفات تحتوي على تفاصيل كافية تسمح بتصميم منظومة الإدارة بما تضمنه من العمليات الإجرائية المختلفة مع بناء البرمجيات والأجهزة بكل تعقيداتها وتفصيلها حيث تستخدم هذه المواصفات في المراحل النهائية لبناء اختبارات مختلفة للحكم على أداء نظام العمل وتحقيقه للمتطلبات التي تم تحديدها في بداية المشروع.

هذا الانتقال التدريجي من مراحل المتطلبات إلى المواصفات ثم إلى اختبارات الأداء هو ما يميز كفاءة فريق التطوير ودقة المنهجية التي استخدموها، ويمكننا التعرف على المصطلحات الرئيسية التالية المرتبطة بأنشطة تحديد المتطلبات والمواصفات ومستوياتها:

- مواصفات المتطلبات: هي وثيقة تحدد مواصفات ما نسعى الحصول عليه.
- تحليل المتطلبات: هي عمليات وأنشطة التحليل التي نقوم بها لإنتاج وثيقة مواصفات المتطلبات
- تعريف المتطلبات: هي العمليات التي نقوم بها خلال مراحل دورة حياة التطوير وحتى نهاية مرحلة التحليل لإنتاج وثيقة المواصفات بما فيها تحديد المتطلبات ودراسات الجدوى والتحليل وبما تضمنه من أنشطة ومراحل جزئية.

2-2- آلية جمع المتطلبات

تعد هذه المرحلة هي أهم المراحل في عملية التطوير وتبدأ هذه المرحلة بالتعرف على المستخدمين والمعنيين بإنتاج البرنامج والتعرف على متطلباتهم بشكل دقيق وواضح، وهناك عدة تقنيات وطرق تساعد في عملية جمع المتطلبات مثل:

- 1 . المقابلة: لقد استخدمنا هذه الطريقة مع أهم الأشخاص المعنيين بإنتاج البرنامج مثل المحللين الفنيين
- 2 . تقنية ورش العمل: وفي هذه الطريقة يقوم كفريق جمع المتطلبات بجمع جميع المعنيين بإنتاج البرنامج وجميع الأشخاص الذين قد يؤثر البرنامج على عملهم في مكان واحد وذلك للاستماع إلى جميع الآراء
- 3 . تقنية العصف الذهني: استخدمنا هذه التقنية علينا كفريق العمل وأيضاً على المستخدمين، وتم في هذه التقنية اقتراح أفكار حول البرنامج
- 4 . استخدام المجسمات والتصاميم الاختبارية: استخدمنا أيضاً هذه التقنية لإطلاع الزبون على ما وصل إليه العمل وأخذ آراءه وكذلك لأخذ الموافقة من الزبون على أن هذه المتطلبات صحيحة

2-3- المتطلبات الوظيفية

تتعلق هذه المتطلبات بالاجرائيات التي يجب على النظام أن ينفذها لتحقيق هدفه والخدمات التي يقدمها للمستخدم والوظائف المختلفة.

يجب أن يقدم النظام لمستخدم البرنامج:

- 1- إمكانية تحصيل مجموعات البيانات الكبيرة
- 2- تخزين البيانات بحيث يمكن إعادة استخدامها
- 3- تطبيق معالجة أولية على البيانات
- 4- إمكانية تمثيل البيانات رسومياً
- 5- تقسيم البيانات المتوفرة لمجموعتين للتدريب والاختبار
- 6- تطبيق نموذج تنبؤ
- 7- إمكانية عرض النتائج لنموذج التنبؤ
- 8- تمثيل نتائج التنبؤ رسومياً مقارنة بالنتائج الفعلية
- 9- حساب معدل الخطأ ودقة النموذج
- 10- إمكانية إجراء عملية التنبؤ
- 11- عرض نتائج عملية التنبؤ ودقة النموذج
- 12- عرض التمثيل الرسومي للبيانات

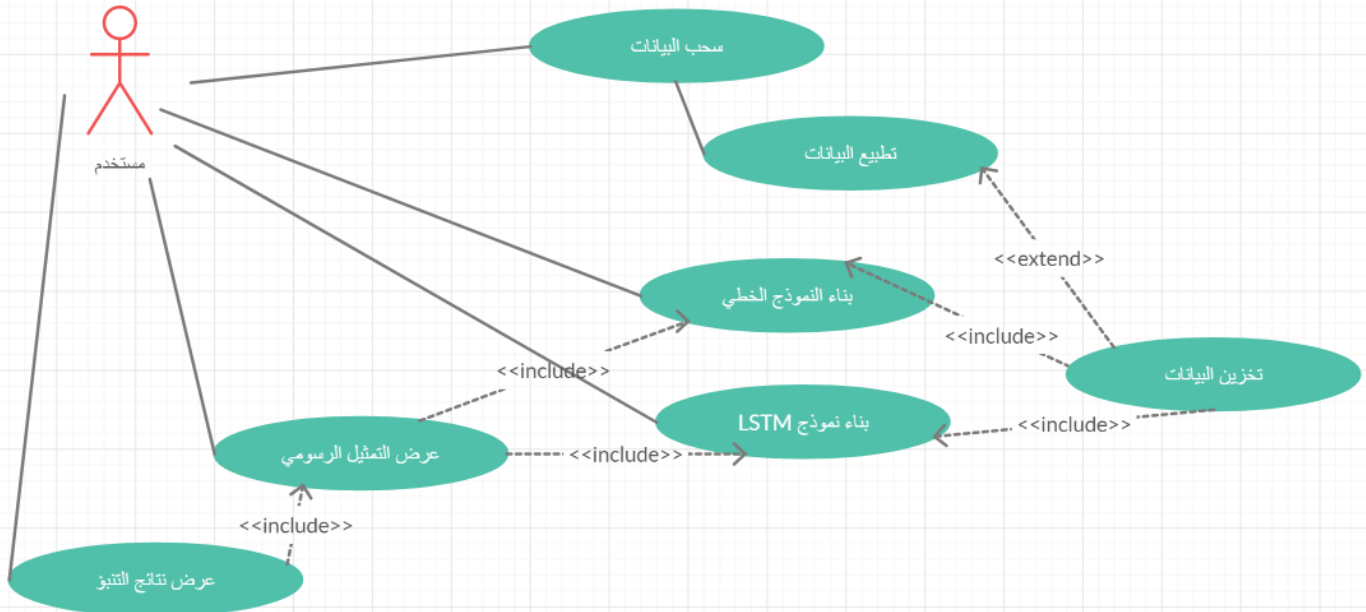
4-2- المتطلبات غير الوظيفية

تتعلق هذه المتطلبات بالموصفات التي يجب أن يتميز بها النظام من حيث خصائصه التي يوفرها للمستخدمين وتحقيق أعلى جودة ممكنة كالفعالية والدقة وسهولة الاستخدام والخصوصية والأمان ومن المتطلبات التي يجب أن يوفرها النظام المراد تطويره تتضمن المتطلبات غير الوظيفية:

- 1- أجهزة حاسوب تتوفر على نظام تشغيل Windows 7,8,10
- 2- اتصال انترنت لتحصيل البيانات
- 3- وجود دعم للغة البرمجة Python على حاسوب العميل
- 4- توفر المكتبات الداعمة على حاسوب العميل

5-2- مخطط حالات الاستخدام Use Case

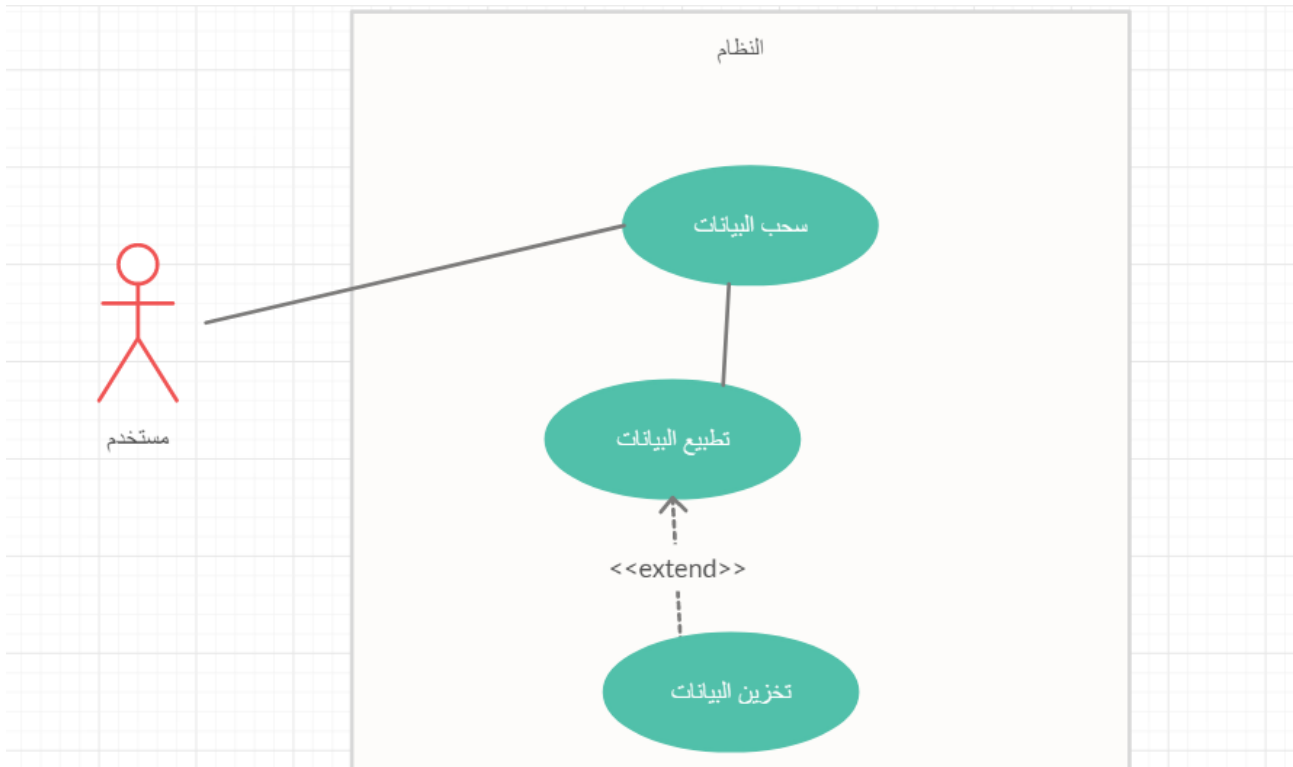
المخطط العام الخاص بالمستخدم



الشكل (9) مخطط حالة الاستخدام للمستخدم

سنقوم الآن بشرح بعض الجزئيات الموجودة:

1- تحصيل (سحب) البيانات

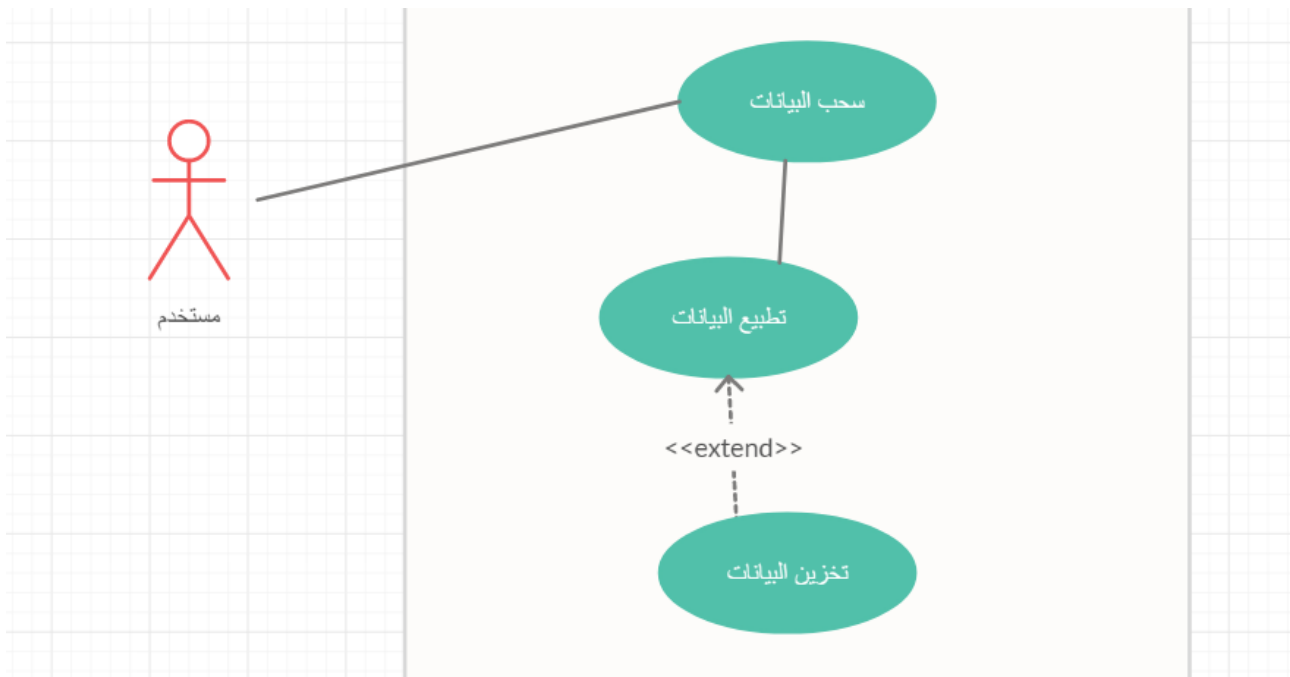


الشكل (10) مخطط حالة الاستخدام لسحب البيانات

معرف حالة الاستخدام UC_1		
اسم حالة الاستخدام	سحب البيانات Data Collect	
الفاعلين	المستخدم	
الحالة المسبقة	تشغيل النظام	
الوصف	يتم فيها تحصيل البيانات المتعلقة بأسهم الشركة من Yahoo Finance	
السيناريو الرئيسي	الفاعل (Action)	رد الفعل (Reaction)
	<ul style="list-style-type: none"> - يتم تشغيل البرنامج - طلب تحصيل البيانات المطلوبة 	<ul style="list-style-type: none"> - يقوم النظام باستدعاء البيانات عبر API - تخزين البيانات في CSV

الجدول (4) حالة الاستخدام سحب البيانات

2- تطبيع البيانات

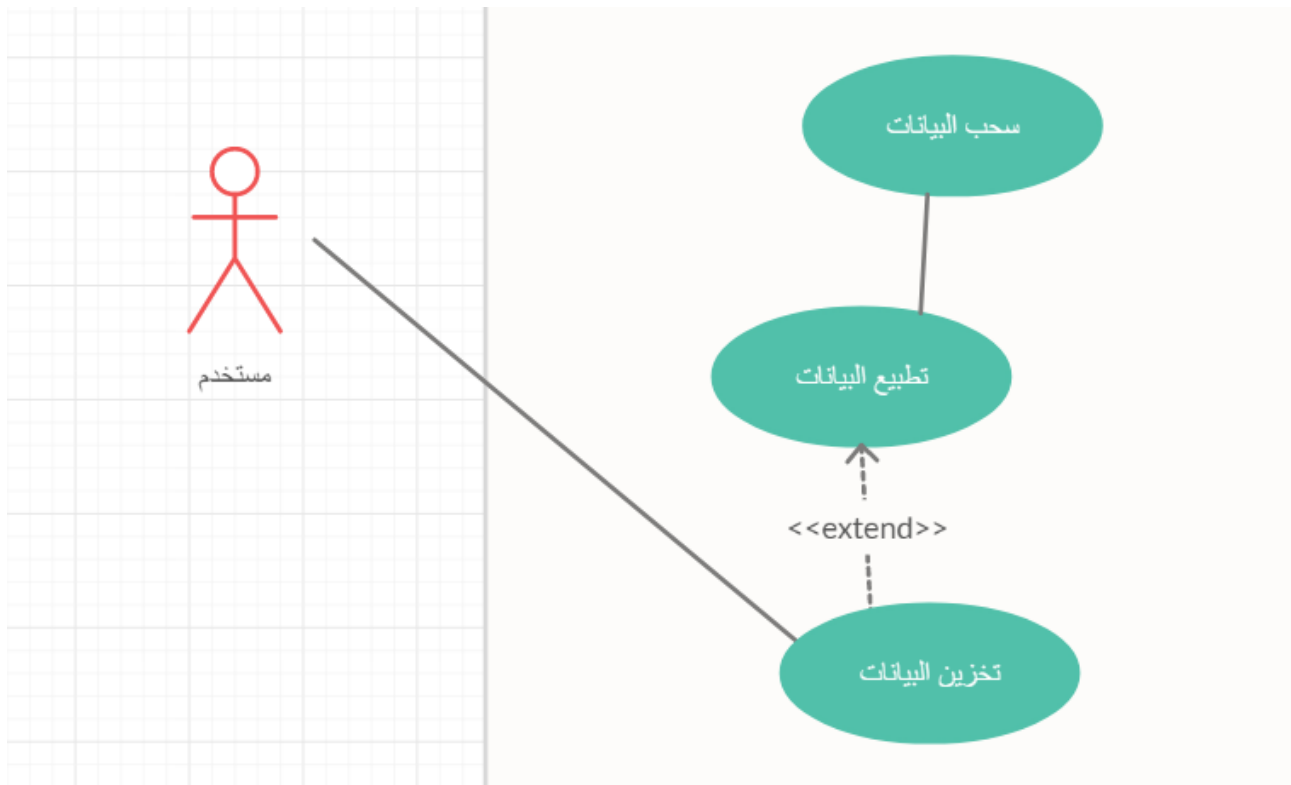


الشكل (11) مخطط حالة الاستخدام لحالة تطبيع البيانات

معرف حالة الاستخدام UC_2		اسم حالة الاستخدام
تطبيع البيانات Normalization		الفاعلين
المستخدم		الحالة المسبقة
سحب البيانات وتخزين البيانات		الوصف
تصف هذه الحالة المرحلة التي يتم فيها تحويل البيانات من قيم رقمية كبيرة إلى قيم بين الصفر والواحد ليسهل تطبيق نموذج تنبؤ عليها بفاعلية		السيناريو الرئيسي
رد الفعل (Reaction)	الفعل (Action)	
- يقوم النظام بتطبيع البيانات - تخزين القيم الجديدة	- يتم سحب البيانات - تخزين البيانات	

الجدول (5) حالة الاستخدام تطبيع البيانات

3- تخزين البيانات

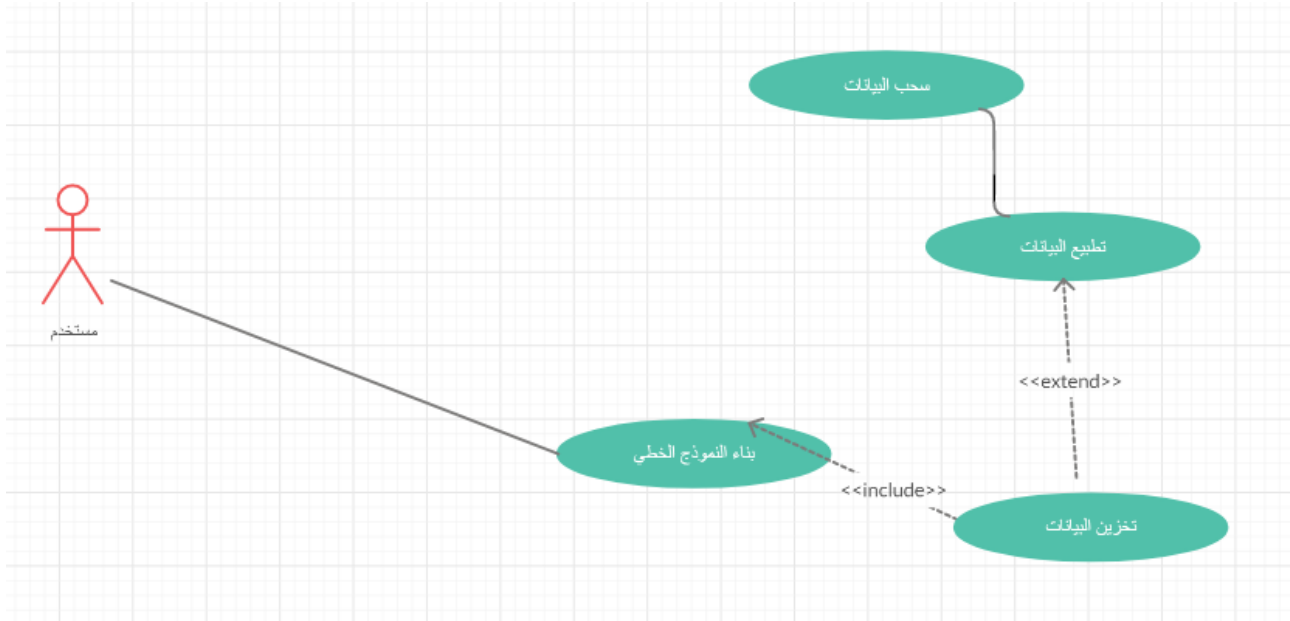


الشكل (12) مخطط حالة الاستخدام لحالة تخزين البيانات

معرف حالة الاستخدام UC_3		
اسم حالة الاستخدام	تخزين البيانات	
الفاعلين	المستخدم	
الحالة المسبقة	تحصيل البيانات	
الوصف	يجري سحب البيانات ومن ثم تخزينها في ملف CSV لمجموعة البيانات الأولية ومن ثم للبيانات ما بعد التطبيع	
السيناريو الرئيسي	الفعل (Action)	رد الفعل (Reaction)
	- سحب البيانات - تطبيع البيانات	- تخزين البيانات في ملف CSV

الجدول (6) حالة الاستخدام تخزين البيانات

4- بناء النموذج الخطي

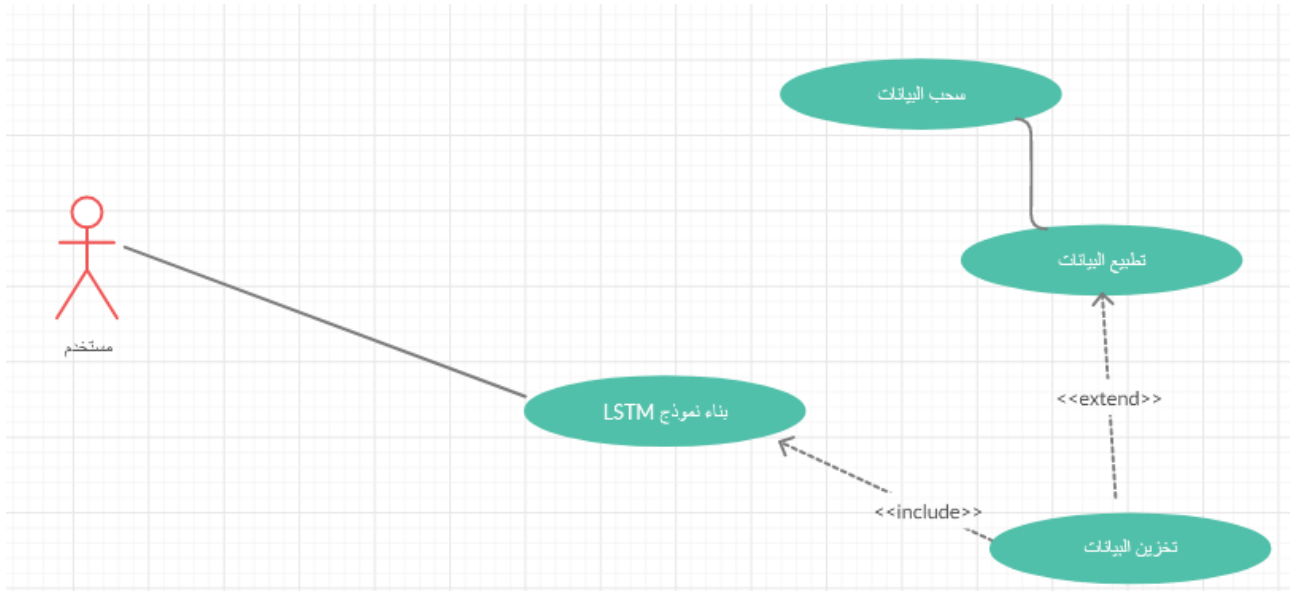


الشكل (13) مخطط حالة الاستخدام لحالة بناء النموذج الخطي

معرف حالة الاستخدام UC_4		
اسم حالة الاستخدام		بناء النموذج الخطي Linear Regression
الفاعلين		المستخدم
الحالة المسبقة		سحب البيانات وتخزين البيانات وتطبيع البيانات
الوصف		يجري بناء النموذج الأولي للتنبؤ باستخدام التراجع الخطي
السيناريو الرئيسي	الفعل (Action)	رد الفعل (Reaction)
	<ul style="list-style-type: none"> - يختار المستخدم - تدريب النموذج 	<ul style="list-style-type: none"> - يجري تدريب النموذج - اجراء التنبؤ - عرض النتائج - عرض التمثيل الرسومي

الجدول (7) حالة الاستخدام بناء النموذج الخطي

5- بناء نموذج LSTM

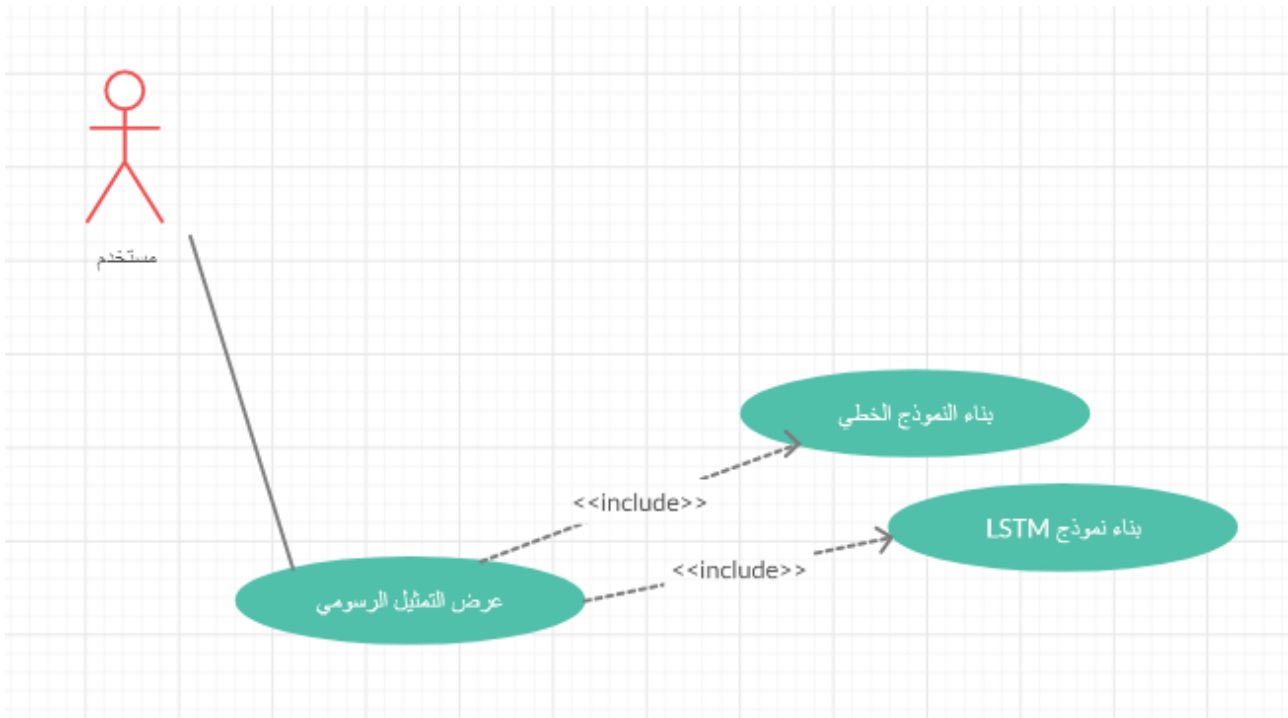


الشكل (13) مخطط حالة الاستخدام لحالة بناء نموذج LSTM

معرف حالة الاستخدام UC_5		
اسم حالة الاستخدام		بناء نموذج التنبؤ باستخدام خوارزمية LSTM
الفاعلين		المستخدم
الحالة المسبقة		سحب البيانات وتخزين البيانات وتطبيع البيانات
الوصف		يجري بناء النموذج الأولي للتنبؤ باستخدام خوارزمية LSTM
السيناريو الرئيسي		<div>الفعل (Action)</div> <div>رد الفعل (Reaction)</div>
<div> <ul style="list-style-type: none"> - يختار المستخدم تدريب النموذج </div>		<div> <ul style="list-style-type: none"> - يجري تدريب النموذج - اجراء التنبؤ - عرض النتائج - عرض التمثيل الرسومي </div>

الجدول (8) حالة الاستخدام بناء نموذج LSTM

6- عرض التمثيل الرسومي

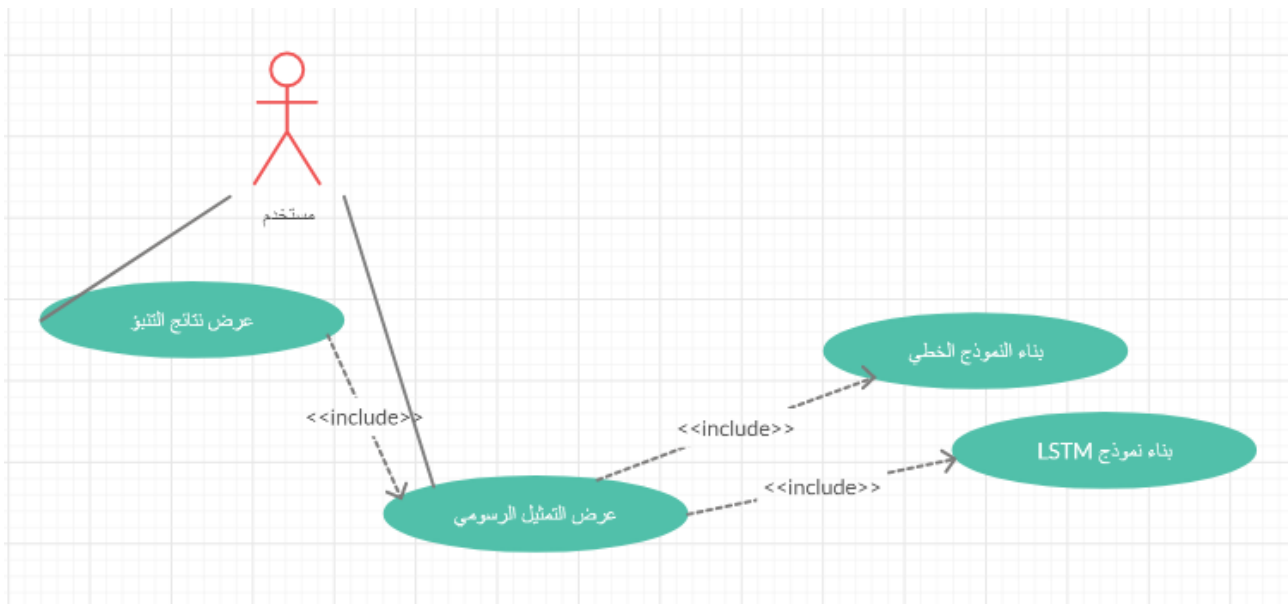


الشكل (14) مخطط حالة الاستخدام لحالة عرض التمثيل الرسومي

معرف حالة الاستخدام UC_6		اسم حالة الاستخدام
عرض التمثيل الرسومي لنتائج التنبؤ		الفاعلين
المستخدم		الحالة المسبقة
بناء نموذج تنبؤ		الوصف
يجري عرض التمثيل الرسومي للبيانات الفعلية مقابل التنبؤ		السيناريو الرئيسي
رد الفعل (Reaction)	الفعل (Action)	
<ul style="list-style-type: none"> - يتم اجراء التدريب - عرض التمثيل الرسومي لنتائج التنبؤ مقابل البيانات الفعلية 	<ul style="list-style-type: none"> - يختار المستخدم تدريب النموذج - يختار عرض التمثيل الرسومي للنتائج 	

الجدول (9) حالة الاستخدام عرض التمثيل الرسومي

7- عرض نتائج التنبؤ وحساب الدقة



الشكل (15) مخطط حالة الاستخدام لحالة عرض نتائج التنبؤ وحساب الدقة

معرف حالة الاستخدام UC_7		
اسم حالة الاستخدام	عرض نتائج التنبؤ وحساب الدقة	
الفاعلين	المستخدم	
الحالة المسبقة	بناء النموذج وتدريبه	
الوصف	بعد تدريب النموذج المختار يتم عرض النتائج وحساب الدقة للتنبؤ باستخدام متوسط مربع الخطأ Mean Squared Error	
السيناريو الرئيسي	الفعل (Action)	رد الفعل (Reaction)
	- يختار المستخدم عرض النتائج	- يتم تدريب النموذج - عرض نتائج التنبؤ - عرض دقة التنبؤ

الجدول (10) حالة الاستخدام عرض نتائج التنبؤ وحساب الدقة

الفصل الرابع

التصميم

إن الهدف من التصميم هو جعل البرمجية التي تم تحليلها قابلة للتطبيق والاستخدام ولتحقيق هذا سوف يجري في هذا الفصل أخذ المعلومات المتعلقة بتقنيات تعلم الآلة والمتطلبات اللازمة التي تم جمعها ونمذجتها في المراحل السابقة وتحويلها إلى كيانات مجردة قابلة للتنفيذ. [1]

1- تقنيات تعلم الآلة

يعرف التعلم الآلي Machine Learning على أنه أتمتة وتحسين عملية تعلم أجهزة الكمبيوتر استناداً إلى تجاربها دون أن تتم برمجتها فعلياً، ومن دون أي مساعدة بشرية حيث تبدأ العملية بتغذية بيانات جيدة ومن ثم تدريب الحاسب من خلال بناء نماذج تعلم الآلة باستخدام البيانات والخوارزميات المختلفة، ويعتمد اختيار الخوارزميات على نوع البيانات التي لدينا ونوع المهمة التي نحاول تنفيذها تلقائياً. وما تزال تقنيات تعلم الآلة في نمو مستمر وهناك الكثير من الأمثلة التي يتم اعتمادها فيها:

- تنبؤ احتمالية عودة مريض للمستشفى خلال مدة معينة
 - توزيع الزبائن لمشروع ما في مجموعات استهلاكية بناء على صفات مشتركة أو مشتريات سابقة متشابهة أو سلوك معين للاستفادة منها في التسويق المستهدف
 - تنبؤ تفاعل مجموعة من الزبائن مع عرض معين
 - تنبؤ بمعدل حالات الإصابة بمرض ما في مجتمع ما (مثل الوباء الحالي)
 - التنبؤ بأسعار الأسهم في سوق التداول (كما هو الحال في هذا المشروع)
- وفي جوهرها كل هذه الأمثلة تهدف للتعلم من البيانات المتاحة والتعامل مع كل حالة يمكن استخدام مجموعة بيانات لتدريب الخوارزمية واستخلاص المعلومات المفيدة منها وهذه الخوارزميات أو تطبيقات التعلم الآلي يمكن تصنيفها في ثلاث فئات رئيسية ، اعتماداً على درجة الحاجة للإشراف والمتابعة خلال التدريب ويعتمد اختيار مصمم النظام على المهمة المتوقعة من النظام، والتصنيفات المقصودة هي على النحو التالي:

1-1- التعلم الخاضع للإشراف Supervised Learning

وهو ما يعتبر نموذجاً للتنبؤ ويستخدم للمهام التي تتطلب تنبؤ بخرج ما output أو بهدف معين باستخدام المتغيرات في مجموعة البيانات dataset وهي عملية تطوير نموذج رياضي أو أداة رياضية لتوليد تنبؤات دقيقة أي عندما تتعلم الخوارزمية من خلال أمثلة البيانات والردود المستهدفة المرتبطة التي يمكن أن تتكون من قيم رقمية أو تسميات مسلسلة مثل الفئات أو العلامات ، من أجل التنبؤ لاحقاً بالاستجابة الصحيحة عند طرحها بأمثلة جديدة ، فإنها تندرج تحت فئة التعلم المشرف، هذا النهج يشبه بالفعل التعلم البشري تحت إشراف المعلم حيث يقدم المعلم أمثلة جيدة للطالب ليحفظها، ثم يستمد الطالب القواعد العامة من هذه الأمثلة المحددة، وهو النمط الذي سوف يتم استخدامه في بناء الخوارزمية الرئيسية لهذا المشروع.

1-2- التعلم غير الخاضع للإشراف Unsupervised Learning

عندما تتعلم الخوارزمية من أمثلة واضحة دون أي استجابة مرتبطة، تترك الخوارزمية لتحديد أنماط البيانات من تلقاء نفسها، ويميل هذا النوع من الخوارزمية إلى إعادة هيكلة البيانات إلى شيء آخر، مثل الميزات الجديدة التي قد تمثل فئة أو سلسلة جديدة من القيم غير المرتبطة، وتعتبر مفيدة جداً في تزويد البشر برؤى ثاقبة لمعنى البيانات والمدخلات المفيدة الجديدة لخوارزميات التعلم الآلي الخاضعة للإشراف.

وباعتبارها من التعلم فهي تشبه الطرق التي يستخدمها البشر لمعرفة أن بعض الكائنات أو الأحداث هي من نفس الفصل، مثل ملاحظة درجة التشابه بين الأشياء وهو ما يسمى بالتعميم أو التصنيف، كأن يتعلم طفل شكل تفاحة ما ويدرك أن كل التفاحات باختلاف أشكالها تبقى تفاحة، وقد تعتمد بعض أنظمة التوصيات الموجودة على الويب في شكل أئمة التسويق الإعلان وتوصية إعلانات معينة لزائر معين على هذا النوع من التعلم.

1-3- التعلم المعزز Reinforcement learning

عندما تقدم الخوارزمية مع أمثلة تفتقر إلى التصنيفات كما هو الحال في التعلم غير الخاضع للإشراف، ومع ذلك يمكن إضافة أمثلة بردود فعل إيجابية أو سلبية وفقاً للحل الذي تقترحه الخوارزمية التي تدرج تحت فئة تعلم التعزيز، وهي متصلة بالتطبيقات التي يجب أن تتخذ الخوارزمية قرارات بشأنها (بحيث يكون المنتج إلزامياً وليس مجرد وصفي كما هو الحال في التعلم غير الخاضع للإشراف)، والقرارات تحمل عواقب في العالم البشري بما يشبه التعلم بالتجربة والخطأ.

تساعدك الأخطاء على التعلم نظراً لوجود عقوبة إضافية (التكلفة وفقدان الوقت والندم والألم وما إلى ذلك) مما يعلمك أن احتمال نجاح مسار معين قد يكون أكثر من الآخرين، وأحد الأمثلة المثيرة على تعلم التعزيز يحدث عندما تتعلم أجهزة الكمبيوتر تشغيل ألعاب الفيديو بمفردها.

في هذه الحالة يقدم التطبيق الخوارزمية مع أمثلة لمواقف معينة، مثل جعل اللاعب عالقاً في متاهة بالإضافة إلى ضرورة تجنب عدو، حيث يتيح التطبيق للخوارزمية معرفة نتائج الإجراءات التي يتخذها، ويحدث التعلم أثناء محاولة تجنب ما يكتشف أنه خطير والاستمرار بمحاولة النجاة، يمكن إلقاء نظرة على كيفية قيام شركة Google DeepMind بإنشاء برنامج تعليمي معزز يلعب ألعاب فيديو Atari القديمة وعند مشاهدة الفيديو يلاحظ كيف يكون البرنامج في البداية غير متقن وغير ماهر ولكن يتحسن بثبات مع التدريب حتى يصبح بطلاً في اللعبة. [6]

2- مفاهيم مرتبطة بتقنيات تعلم الآلة**• البيانات**

يمكن أن تكون أي حقيقة أو قيمة أو نص أو صوت أو صورة غير معالج لا يتم تفسيرها وتحليلها، والبيانات هي الجزء الأكثر أهمية في جميع تحليلات البيانات والتعلم الآلي والذكاء الاصطناعي لأنه بدون بيانات لا يمكننا

تدريب أي نموذج وستذهب كل الأبحاث والأتمتة الحديثة دون جدوى، ما يدفع الشركات الكبيرة لإنفاق الكثير من الأموال فقط لجمع أكبر قدر ممكن من البيانات وهو ما يعرف اليوم بالبيانات الكبيرة Big Data

• المعلومات

البيانات التي تم تفسيرها والتلاعب بها واستنتاج بعض الاستدلالات ذات المغزى بالنسبة للمستخدمين بعد معالجتها.

• المعرفة

مزيج من المعلومات المستخلصة والخبرات والتعلم ويمكن تسميتها بالبصيرة الناتجة في النهاية وهي تعتبر الهدف الذي يسعى لتحقيقه الباحث في البيانات.

3- استكشاف البيانات

في هذه الخطوة سوف نبدأ بشرح البيانات التي تم استخدامها في هذا المشروع وكيفية ارتباطها بمراحل العمل والبيانات التي تم استخدامها في هذا المشروع هي بيانات شركة ألبايت Alphabet Inc ابتداء من 2000-01-01 وحتى 2020/08/01 وهذه القيم لمجموعة من نقاط البيانات المفهرسة زمنياً والمرتبة (أو سلسلة زمنية). والهدف هنا هو التنبؤ بسعر الإغلاق في أي تاريخ مستقبلي معطى بعد إنجاز التدريب على البيانات المتاحة، ومن أجل سهولة تحقيق إعادة الإنتاج والاستخدام، كل البيانات سوف يتم استخراجها باستخدام API من Yahoo finance وباستخدام لغة Python.

التنبؤ سوف يركز على سعر الإغلاق (Adjusted Closing) للأسهم في البيانات المعطاة.

ويكون شكل البيانات التي يتم الحصول عليها موضحاً في الجدول (11)

Date	High	Low	Open	Close	Volume	Adj Close
2000/01/03	1478	1438.36	1469.25	1455.22	9.32E+08	1455.22
2000/01/04	1455.22	1397.43	1455.22	1399.42	1.01E+09	1399.42
2000/01/05	1413.27	1377.68	1399.42	1402.11	1.09E+09	1402.11
2000/01/06	1411.9	1392.1	1402.11	1403.45	1.09E+09	1403.45
...
...
2020/08/03	3302.73	3284.53	3288.26	3294.61	4.64E+09	3294.61
2020/08/04	3306.84	3286.37	3289.92	3306.51	4.62E+09	3306.51
2020/08/05	3330.77	3317.37	3317.37	3327.77	4.73E+09	3327.77
2020/08/06	3351.03	3318.14	3323.17	3349.16	4.27E+09	3349.16
2020/08/07	3352.54	3328.72	3340.05	3351.28	4.1E+09	3351.28

الجدول (11) البيانات الأولية التي يتم تحصيلها

كافة البيانات موجودة في ملف yahoo500.csv ولم يتم ملاحظة وجود أي أخطاء في البيانات حيث لا توجد قيم سالبة أو غير رقمية.

ومن مجموعة البيانات هذه يمكننا استنتاج أن قيم التاريخ والقيمة الأعلى (High) والقيمة الأقل (low) ليست مزايا مهمة للبيانات ولا يهم ما هي أعلى قيمة للبيانات في يوم ما أو أقل قيمة في أحد أيام التداول وما يهم بالنسبة لنا (وفق الشمعات اليابانية) هو سعر الافتتاح للأسهم وسعر الإغلاق.

إذا كان السعر في نهاية اليوم عند الإغلاق أعلى من السعر في بداية اليوم سعر الفتح عند بداية التداول هذا يعني أننا قد حققنا أرباحاً وغير ذلك فهذا يعني خسائراً، بالإضافة إلى أن معلومة حجم التداول في السوق Volume Share ذات أهمية فهذا يعني أن السوق المتجه صعوداً يجب أن يقابله صعود في حجم التداول في السوق، وعلى سبيل المثال: فإن تزايد السعر ولكن انخفاض حجم التداول دليل على فقدان الاهتمام من قبل المستثمرين وهذه معلومة تحذيرية لاحتمال انهيار في سعر الأسهم، حيث إن أي تحرك في سعر الأسهم بزيادة أو نقصان يرافقه ارتفاع في حجم التداول في السوق يعتبر إشارة قوية على وجود تغير حقيقي في السوق، لذلك تم إزالة قيم التاريخ والأعلى والأقل من مجموعة البيانات في مرحلة المعالجة الأولية للبيانات وتم حساب المتوسط والانحراف المعياري والقيمة العظمى والقيمة الدنيا وكانت النتائج:

```

Open --- mean : 1612.8617081157508 Std: 623.7806557037948 Max: 3380.4499509999996
Min: 679.2800292999999
High --- mean : 1622.2491352942918 Std: 625.4272480150993 Max: 3393.5200200000004
Min: 695.2700195
Low --- mean : 1602.6339384709174 Std: 621.8154703306687 Max: 3378.830078 Min:
666.789978
Close --- mean : 1613.0281263259676 Std: 623.840630715595 Max: 3386.149902 Min:
676.5300292999999
Volume --- mean : 3152598171.4396286 Std: 1520150153.7959082 Max: 11456230000 Min:
356070000

```

الشكل (16) نتائج حساب الانحراف المعياري والوسيط للبيانات

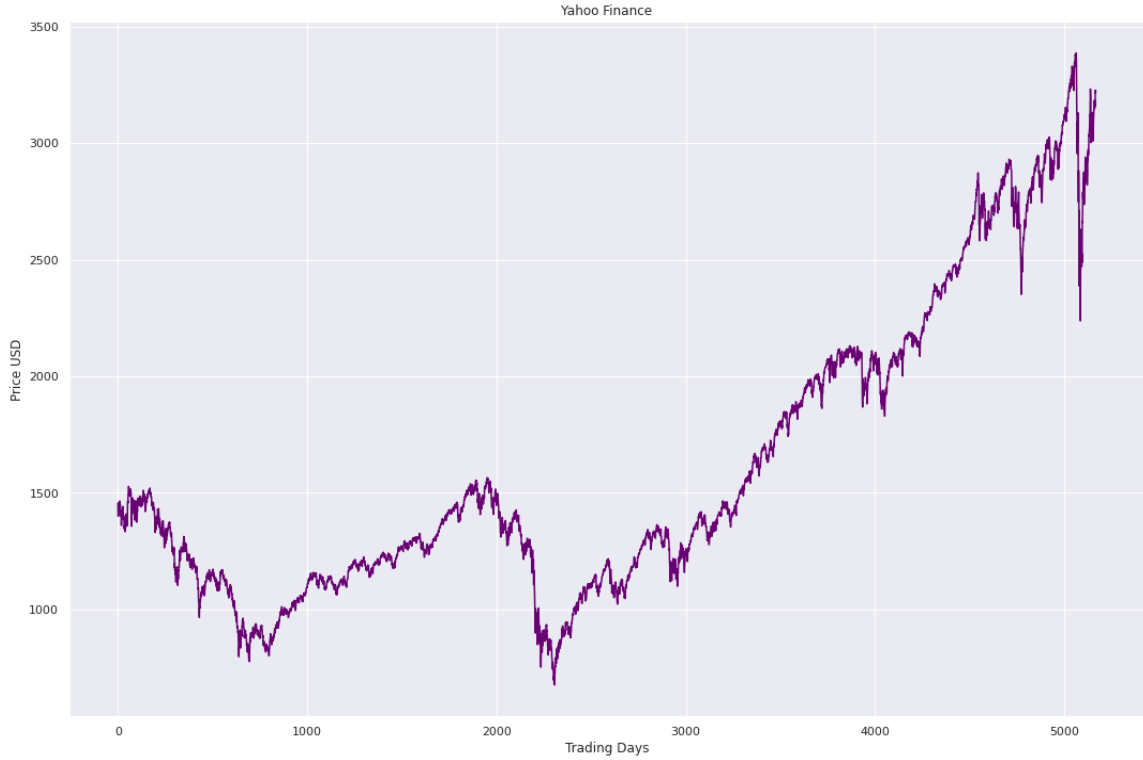
3-1- جدول قواميس البيانات

الوصف		الكيان الذي يحوي المعلومات
Date	التاريخ	datetime
High	أعلى سعر تم تحقيقه في اليوم	float
Low	أقل سعر تم تحقيقه في اليوم	Float
Open	سعر الافتتاح	Float
Close	سعر الإغلاق	Float
Volume	حجم التداول في السوق	Float
Adj Close	سعر الإغلاق الفعلي	Float
Std	الانحراف المعياري	Float
Mean	المتوسط	Flot

الجدول (12) قواميس البيانات

3-2- استكشاف البيانات رسومياً

للتمثيل الرسومي للبيانات تم استخدام مكتبة matplotlib وتم رسم النقاط المختلفة لأسعار إغلاق الأسهم للبيانات ويرافقها عدد الأيام المتاحة ويلاحظ بأن عدد العينات تزيد عن 5000 عينة تمثل الأيام التي تم تضمينها في مجموعة البيانات:



الشكل (17) تمثيل البيانات رسوماً

X تمثل عدد العينات أي سعر الإغلاق يومياً Y تمثل السعر بالدولار الأمريكي

رغم أن البيانات تمثل اتجاهات تصاعدياً يعبر عن تزايد السعر باستمرار (لشركة ألفابت) يمكن ملاحظة انهيار كبير يعبر عن الأزمة المالية العالمية في 2008 (بين الأيام 2000-3000) وملاحظة انهيار إضافي (عند الأيام 5000) غالباً يعبر عن مرحلة الوباء العالمي COVID19

4- الخوارزميات والتقنيات

هدف هذا المشروع هو دراسة بيانات من السلاسل الزمنية واستكشاف أكبر قدر من ممكن من الخيارات المتاحة للتنبؤ الدقيق بأسعار الأسهم، ومن خلال البحث وجدنا مواضيع عديدة متعلقة بالشبكات العصبونية المتكررة (Recurrent Neural Nets (RNN) والمستخدمه خصوصاً للتسلسل والتعلم بالأنماط حيث تعبر عن شبكات وحلقات تكرارية بداخلها، مما يساعد في الحفاظ على المعلومات بالتالي القدرة في حفظ البيانات بشكل دقيق. ولكن هذه الشبكات تعاني من مشكلة في التعلم من البيانات التاريخية كما هو مفترض وتم حل هذه المشكلة في خوارزمية الذاكرة طويلة قصيرة المدى Long-Short Term Memory Networks واختصاراً LSTM وتعبر هذ الخوارزمية عن نوع خاص من شبكات RNN قابلة للتعلم على المدى الطويل. وبالإضافة للاختلاف المعياري عن الشبكات العصبونية، فإن هذه المقاييس الإضافية تعتبر مفيدة في تحسين نموذج التنبؤ:

- الإدخال: معالجة أولية للبيانات وتطبيع

- بنية الشبكات العصبونية :

- عدد الطبقات (كم طبقة في النموذج وتم استخدام ثلاثة)
- عدد العقد (عدد العقد في الطبقة)

- مقاييس التدريب

- التدريب : توزيع البيانات بين تدريب واختبار (82.95% و 17.05%)
- مجموعات التحقق : تشكل 0.05%
- حجم الدفعة Batch Size : تمثل عدد الخطوات التي يجب تضمينها في التدريب الواحد وتم استخدام 1 للنموذج الأساسي في lstm و 512 في النموذج المحسن
- دالة التحسين
- المراحل (Epoch) : عدد المرات التي يجب تنفيذ عملية التدريب فيها وتم استخدام 1 للنموذج الأساسي و 20 في النموذج المحسن.

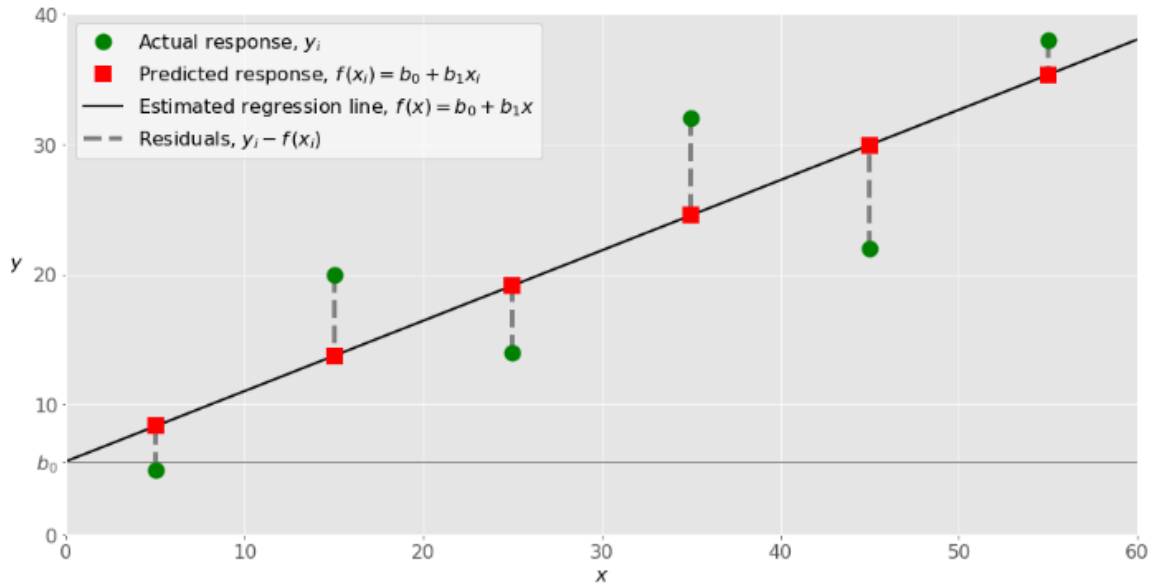
5- النموذج المعياري وفق التراجع الخطي Linear Regression Model

بهدف تحقيق هذا المشروع تم استخدام نموذج التراجع الخطي كنموذج معياري مبدئي حيث أحد أهداف هذا المشروع هي فهم التطبيقات المختلفة لآليات تعلم الآلة. يتميز التراجع الخطي بكونه نموذجاً بسيطاً ويمثل معادلة خطية تجمع بين مجموعة محددة من قيم الدخل ولتكن X مع الحلول للخروج المتنبأ به ولتكن Y ويمكن تمثيلها بالمعادلة:

$$Y = B0 + B1 * X$$

حيث B0 و B1 معاملات تدخل في النموذج

ومثالاً عليه الشكل (18)



الشكل (18) مثال على تراجع خطي بسيط

6- المنهجية

سوف نتطرق لشرح منهجية التصميم المتبعة في إنجاز المشروع:

6-1- المعالجة الأولية للبيانات

تحصيل البيانات وإجراء المعالجة الأولية عليها لإنجاز المشروع تحدث وفق التسلسل التالي والتي تم دمج الاجرائيات المتعلقة بها في ملف واحد تحت اسم preprocess_data.py ليتم استدعائه في البرنامج التنفيذي:

6-1-1 تحصيل البيانات:

طلب البيانات من موقع ياهو للتداول وأسعار الأسهم yahoo finance باستخدام API ثم حفظ البيانات المحصّلة في ملف CSV تحت اسم yahoosp500.csv وفق الصيغة التالية:

	Date	Open	High	Low	Close	Volume
	2020/07/17	3224.21	3233.52	3205.65	3224.73	3.99E+09
	2020/07/16	3208.36	3220.39	3198.59	3215.57	3.96E+09
	2020/07/15	3225.98	3238.28	3200.76	3226.56	4.67E+09
	2020/07/14	3141.11	3200.95	3127.66	3197.52	4.48E+09
	2020/07/13	3205.08	3235.32	3149.43	3155.22	4.89E+09
	2020/07/10	3152.47	3186.82	3136.22	3185.04	4.52E+09
	2020/07/09	3176.17	3179.78	3115.7	3152.05	4.83E+09

الجدول (13) يمثل شكل البيانات الأولية المحصّلة

2-1-6 إزالة البيانات غير المهمة

إزالة الخصائص الغير مهمة في عملية التنبؤ (مثل : التاريخ date، أعلى قيمة high ، أقل قيمة low) من البيانات المحصلة وتصبح النتيجة كما يلي :

Item	Open	Close	Volume
0	1469.250000	1455.219971	931800000
1	1455.219971	1399.420044	1009000000
2	1399.420044	1402.109985	1085500000
3	1402.109985	1403.449951	1092300000
4	1403.449951	1441.469971	1225200000
..
..
5163	3205.080078	3155.219971	4890780000
5164	3141.110107	3197.520020	4476170000
5165	3225.979980	3226.560059	4669760000
5166	3208.360107	3215.570068	3961230000

الجدول (14) البيانات بعد إزالة الخصائص غير المهمة

3-1-6 تطبيع البيانات

عوضاً عن التعامل مع قيم رقمية كبيرة سوف يتم تطبيع البيانات وتحويل الأرقام المتعلقة بقيمة سعر الفتح والإغلاق وحجم السوق إلى أرقام ما بين صفر وواحد باستخدام الطريقة الإحصائية والموجود في مكتبة Scikit-learn بدالة اسمها MinMaxScaler لتصبح النتيجة كما يلي :

0	0.292455	0.28738	0.051867
1	0.287261	0.266787	0.058822
2	0.266603	0.267779	0.065713
3	0.267599	0.268274	0.066326
4	0.268095	0.282305	0.078299
5	0.28217	0.288258	0.063849
6	0.288142	0.281231	0.059272
...
5160	0.915822	0.920207	0.411853
5161	0.924373	0.913604	0.402963
5162	0.9156	0.925779	0.374704
5163	0.935076	0.914774	0.408527
5164	0.911394	0.930385	0.371175

5165	0.942814	0.941102	0.388615
5166	0.936291	0.937047	0.324785

الجدول (15) جدول يمثل شكل البيانات بعد التطبيع

تم تخزين البيانات بعد التطبيع في ملف yahoo_preprocessed.csv من أجل إعادة الاستخدام المستقبلية.

6-2- تدريب البيانات للنموذج المعياري باستخدام التراجع الخطي

تقسيم مجموعة البيانات dataset إلى مجموعتين الأولى للتدريب وتشكل (68.53%) والثانية للاختبار وتشكل (31.47%) من أجل نموذج التراجع الخطي. وكان شكل تقسيم البيانات كما يلي :

```
x_train (3541, 1)
y_train (3541, 1)
x_test (1627, 1)
y_test (1627, 1)
```

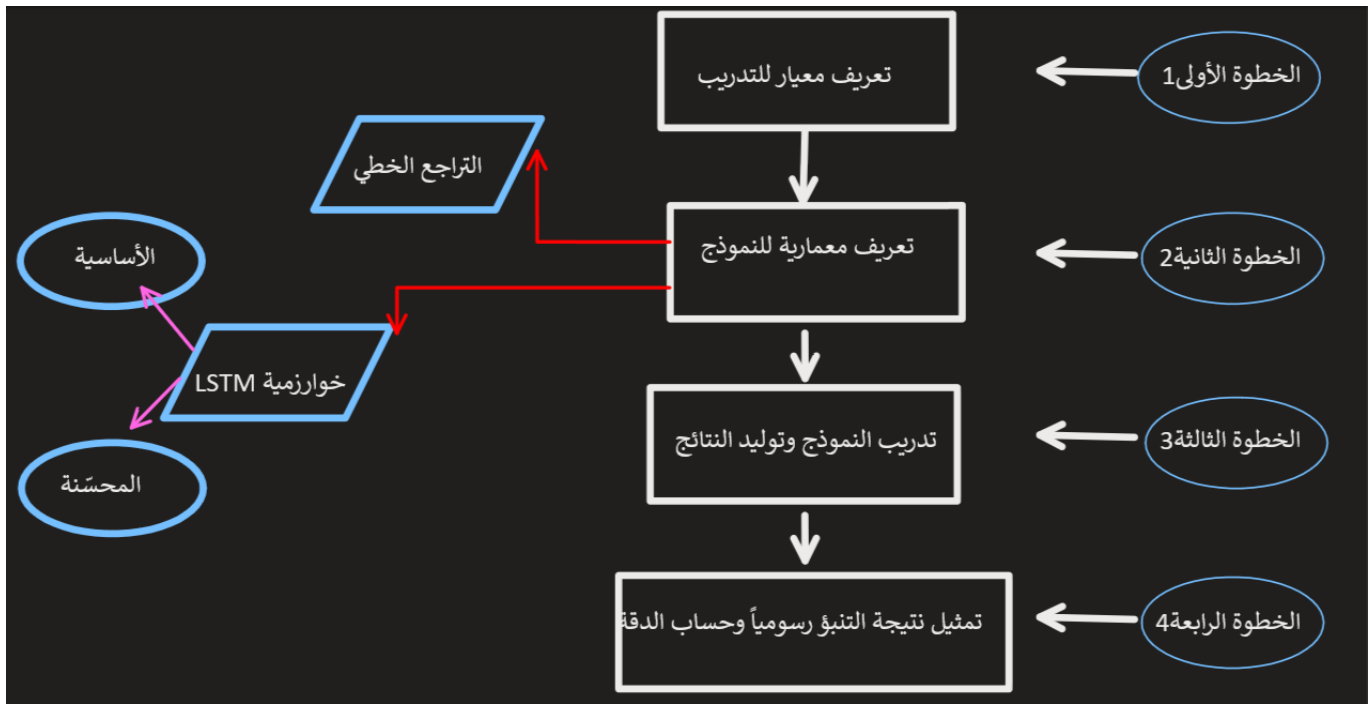
6-3- تدريب البيانات لتطبيق الخوارزمية النهائية LSTM

تم تقسيم مجموعة البيانات dataset إلى مجموعتين الأولى للتدريب وتشكل (82.95%) والثانية للاختبار وتشكل (17.05%) من أجل نموذج LSTM. وكان شكل تقسيم البيانات كما يلي :

```
x_train (4612, 50, 3)
y_train (4612,)
x_test (446, 50, 3)
y_test (446,)
```

7- التنفيذ

بعد أن تم تنزيل البيانات وتطبيق المعالجة الأولية عليها سوف يتم إنجاز عملية التنفيذ بشكل وفق النماذج المرحلية التالية كما يلي :



الشكل (19) مراحل تنفيذ النظام

الفصل الخامس

التحقيق البرمجي

تم تنفيذ التطبيق على جداول البيانات التي تم تحصيلها وتخزينها محلياً [2]

1- إجراءات المعالجة

تم تعريف الإجراءات المخزنة التالية في ملف preprocess_data.py :

- الإجراءية get_normalised_data(data)

```
def get_normalised_data(data):
```

تقوم بتطبيع قيم البيانات باستخدام MinMaxScaler من مكتبة sklearn، ويمرر لها DataFrame بالأعمدة التالية ['index', 'Open', 'Close', 'Volume'] وتعيد DataFrame يحوي قيم مطبّعة لكل الأعمدة ما عدا الفهرسة index

- الإجراءية remove_data (data)

```
def remove_data(data):
```

تقوم بإزالة الأعمدة الغير مرغوبة من مجموعة البيانات، ويمرر لها DataFrame بالأعمدة التالية ['Date','Open','High','Low','Close','Volume'] وتعيد DataFrame بالأعمدة التالية بعد الحذف وإضافة فهرسة ['index','Open','Close','Volume']

تم تعريف الإجراءات المخزنة التالية في ملف LinearRegressionModel.py :

- الإجراءية build_model(X, y):

```
def build_model(X, y):
```

تقوم ببناء نموذج للتراجع الخطي باستخدام sklearn.linear_model ويمرر لها في X البيانات المختارة و Y عنوان البيانات

- الإجراءية predict_prices

```
def predict_prices(model, x, label_range):
```

تقوم بالتنبؤ بالأسعار لعنوان معين من البيانات ويمرر لها النموذج وبيانات الاختبار ومجال البيانات المطبّعة

تم تعريف الإجراءات المخزنة التالية في ملف stock_data.py :

- الإجراءية scale_range(x, input_range, target_range)

```
def scale_range(x, input_range, target_range):
```

تقوم هذه الإجراءية بإعادة تقييس مصفوفة Numpy من الدخل إلى مجال الهدف ويمرر لها X البيانات التي يجب إعادة تقييسها، ومجال الدخل، ومجال الهدف وتعيد مصفوفة معاد تقييسها بين مجالين [min, max]

- الإجراءية train_test_split_linear_regression(stocks)

```
def train_test_split_linear_regression(stocks):
```

تقوم بتقسيم مجموعة البيانات dataset إلى بيانات للتدريب وبيانات للاختبار من أجل بناء نموذج التراجع الخطي ويمرر مجموعة البيانات كاملة تحوي الخصائص ['Open','Close','Volume']

- الإجراءية train_test_split_lstm

```
def train_test_split_lstm(stocks, prediction_time=1, test_data_size=450, unroll_length=50):
```

تقوم بتقسيم مجموعة البيانات إلى قسم للتدريب وقسم للاختبار من أجل بناء نموذج الذاكرة قصيرة – طويلة المدى (LSTM)

تم تعريف الإجراءات المخزنة التالية في ملف lstm.py :

- الإجراءية build_basic_model(input_dim, output_dim, return_sequences)

```
def build_basic_model(input_dim, output_dim, return_sequences):
```

تقوم ببناء نموذج LSTM أولي يمرر لها أبعاد الدخل وأبعاد الخرج وتعيد نموذج مبدئي بثلاث طبقات layers

- الإجراءية build_improved_model(input_dim, output_dim, return_sequences)

```
def build_improved_model(input_dim, output_dim, return_sequences):
```

تقوم ببناء نموذج LSTM محسّن باستخدام `keras.layers.recurrent.lstm` يمرر لها أبعاد الدخل وأبعاد الخرج وتعيد نموذج بثلاث طبقات `layers`

تم تعريف الإجراءات المخزنة التالية في ملف `visualize.py` :

• الإجراءية `plot_basic`

```
def plot_basic(stocks, title='Yahoo Finance', y_label='Price USD', x_label='Trading Days'):
```

تقوم بعرض الشكل الرسومي (تمثيل البيانات رسومياً) لأسعار الإغلاق والبيانات الأولية بعد إزالة البيانات غير المهمة.

• الإجراءية `plot_prediction`

```
def plot_prediction(actual, prediction, title='Yahoo Finance vs Prediction', y_label='Price USD', x_label='Trading Days'):
```

تقوم بعرض كل من بيانات التدريب والاختبار والتنبؤ رسومياً (تمثيل البيانات رسومياً) أي القيم المتنبأ بها مقابل قيم الإغلاق الفعلية

• الإجراءية `plot_lstm_prediction`

```
def plot_lstm_prediction(actual, prediction, title='Yahoo Finance vs Prediction', y_label='Price USD', x_label='Trading Days'):
```

تقوم بتمثيل البيانات المتعلقة بالخوارزمية المطبقة LSTM وعرض كل من بيانات التدريب والاختبار والتنبؤ

2- تسلسل الإجراءات

هدف هذا المشروع هو التنبؤ الدقيق بأسعار الإغلاق المستقبلية لأسهم شركات معينة على فترات زمنية مستقبلية. لتحقيق هذه المشروع تم استخدام خوارزمية تحت اسم LSTM للتنبؤ بأسعار أسهم شركة ألفا وأسهم S&P 500 باستخدام بيانات تاريخية [7]

2-1- الخطوة الأولى

تعريف الدالة التي سوف تقوم باستحضار البيانات من قاعدة بيانات ياهو للتداول Yahoo finance

```
sp500 = data.DataReader('^GSPC', data_source='yahoo', start='1/1/2000',
                        end='17/07/2020')

sp500.to_csv('yahoosps500.csv', index=True)
```

بعد حفظ البيانات محلياً في ملف 'yahoosps500.csv' ننتقل إلى الخطوة التالية

2-2- الخطوة الثانية

1-2-2 معالجة أولية للبيانات بالاستفادة من مكتبة Pandas

حساب المتوسط mean والانحراف المعياري standard deviation والقيمة العظمى Max والدنيا Min

```
import pandas as pd

data = pd.read_csv('yahoosp500.csv')
print(data.head())

print("\n")
print("Open    --- mean :", np.mean(data['Open']), " \t Std: ",
      np.std(data['Open']), " \t Max: ", np.max(data['Open']), " \t Min: ",
      np.min(data['Open']))
print("High    --- mean :", np.mean(data['High']), " \t Std: ",
      np.std(data['High']), " \t Max: ", np.max(data['High']), " \t Min: ",
      np.min(data['High']))
print("Low     --- mean :", np.mean(data['Low']), " \t Std: ",
      np.std(data['Low']), " \t Max: ", np.max(data['Low']), " \t Min: ",
      np.min(data['Low']))
print("Close   --- mean :", np.mean(data['Close']), " \t Std: ",
      np.std(data['Close']), " \t Max: ", np.max(data['Close']), " \t Min: ",
      np.min(data['Close']))
print("Volume  --- mean :", np.mean(data['Volume']), " \t Std: ",
      np.std(data['Volume']), " \t Max: ", np.max(data['Volume']), " \t Min: ",
      np.min(data['Volume']))
```

2-2-2 إزالة البيانات غير الضرورية : مثل التاريخ وأكبر قيمة وأدنى تحققت في اليوم

```
import preprocess_data as ppd
stocks = ppd.remove_data(data)
```

3-2-2 تمثيل البيانات الخام رسوماً كما هي في وضعها الحالي

```
visualize.plot_basic(stocks)
```

4-2-2 تطبيع البيانات باستخدام دالة MinMaxScaler لتحويلها إلى قيم بين الصفر والواحد وحفظ البيانات في ملف جديد

```
stocks = ppd.get_normalised_data(stocks)
stocks.to_csv('yahoo_preprocessed.csv', index=False)
```

5-2-2 تمثيل البيانات بعد التطبيع رسوياً

```
visualize.plot_prediction(y_test, predictions)
```

2-3- الخطوة الثالثة

سوف يتم في هذه المرحلة بناء النموذج الأولي بالاعتماد على التراجع الخطي والبيانات التي سوف يتم تطبيق النموذج عليها هي البيانات بعد المعالجة الأولية من المرحلة السابقة.

1-3-2 تقسيم البيانات:

يتم تقسيم البيانات إلى بيانات للتدريب وبيانات للاختبار بهدف تطبيق نموذج التراجع الخطي باستدعاء الدالة (المعرفة في ملف stock_data.py) والتي تقوم بتقسيم البيانات من أجل نموذج التراجع الخطي وفيما يلي تطبيق تنفيذ الدالة:

```
linearstocks = pd.read_csv('yahoo_preprocessed.csv',nrows=1250)
# split the data
X_train, X_test, y_train, y_test, label_range = sd.train_test_split_linear_regression(linearstocks)
```

2-3-2 تدريب النموذج الأولي

في هذه المرحلة يتم بناء النموذج بالاعتماد على مكتبة scikit-learn linear_model واستدعاء الدالة المعرفة في ملف LinearRegressionModel.py والتي تقوم ببناء النموذج الخاص للمشروع كما يلي :

```
# train a linear regresson
model = LinearRegressionModel.build_model(X_train, y_train)
```

3-3-2 الحصول على التنبؤ لمجموعة بيانات الاختبار:

```
# get predict on test set
predictions = LinearRegressionModel.predict_prices(model, X_test, label_range)
```

4-3-2 التمثيل الرسومي لنتائج التنبؤ باستخدام النموذج الأولي باستخدام التراجع الخطي مقارنة بالقيم الفعلية

```
visualize.plot_prediction(y_test, predictions)
```

5-3-2 حساب دقة التنبؤ ونتيجة الاختبار لنتائج النموذج الأولي

لحساب نتيجة التدريب نقوم بحساب متوسط مربع الخطأ Mean Squared Error (MSE) ومن ثم جذر متوسط مربع الخطأ Root Mean Squared Error :

```
trainScore = mean_squared_error(X_train, y_train)
print('Train Score: %.4f MSE (%.4f RMSE)' % (trainScore, math.sqrt(trainScore)))
```

لحساب نتيجة الاختبار نقوم بحساب متوسط مربع الخطأ Mean Squared Error (MSE) ومن ثم جذر متوسط مربع الخطأ Root Mean Squared Error

```
testScore = mean_squared_error(predictions, y_test)
print('Test Score: %.8f MSE (%.8f RMSE)' % (testScore, math.sqrt(testScore)))
```

4-2- الخطوة الرابعة

في هذه الخطوة سوف يتم بناء نموذج التنبؤ بالاعتماد على خوارزمية تعلم الآلة تحت اسم الذاكرة طويلة قصيرة المدى Long-Sort Term Memory اختصاراً (LSTM) وفي هذا الجزء سوف يتم استخدام الخوارزمية للتدريب والاختبار على مجموعة البيانات.

1-4-2 بناء نموذج تنبؤ أولي باستخدام LSTM

بغرض بناء النموذج الأولي لخوارزمية LSTM سوف نقوم بتضمين مكتبة Keras لبناء النموذج وبلاستفادة من مكتبة TensorFlow والبيانات التي سوف يتم العمل عليها ذاتها البيانات بعد المعالجة الأولية.

2-4-2 تقسيم البيانات إلى بيانات للتدريب وبيانات للاختبار واختبار البيانات لنموذج lstm

```
#Split train and test data sets and Unroll train and test data for lstm model
X_train, X_test, y_train, y_test = sd.train_test_split_lstm(stocks_data, 5)

unroll_length = 50
X_train = sd.unroll(X_train, unroll_length)
```



```
X_test = sd.unroll(X_test, unroll_length)
y_train = y_train[-X_train.shape[0]:]
y_test = y_test[-X_test.shape[0]:]
```

3-4-2 بناء نموذج تنبؤ أولي باستخدام الذاكرة طويلة قصيرة المدى

```
# build basic lstm model
model = lstm.build_basic_model(input_dim=X_train.shape[-1], output_dim=unroll_length, return_sequences=True)
```

4-4-2 تدريب النموذج الأولي من خوارزمية LSTM

بالاعتماد على مكتبة TensorFlow سوف يجري تدريب النموذج [3]

```
def train_model(self):
    from Stock_Price_Predictor import model, X_train, y_train
    model.fit(
        X_train,
        y_train,
        batch_size=1,
        epochs=1,
        validation_split=0.05)
```

5-4-2 اجراء عملية التنبؤ باستخدام بيانات الاختبار

```
predictions = model.predict(X_test)
```

6-4-2 تمثيل نتيجة التنبؤ باستخدام النموذج الأولي رسومياً

```
visualize.plot_lstm_prediction(predictions, y_test)
```

7-4-2 حساب دقة التنبؤ ونتيجة الاختبار لنتائج النموذج الأولي من خوارزمية LSTM:

لحساب نتيجة التدريب نقوم بحساب متوسط مربع الخطأ Mean Squared Error (MSE) ومن ثم جذر متوسط

مربع الخطأ Root Mean Squared Error

5-2- الخطوة الخامسة

بناء النموذج المحسن من خوارزمية LSTM:

تهدف هذه الخطوة إلى بناء نموذج تنبؤ بالاعتماد على ذات الخوارزمية ولكن بهدف تحقيق نتائج أفضل وفق الخطوات التالية:

1-5-2 تقسيم البيانات إلى بيانات للتدريب وبيانات للاختبار واختبار البيانات لنموذج lstm

ونلاحظ أنه تم استخدام نفس البيانات بعد المعالجة الأولية والتي تم استخدامها في بناء النموذجين السابقين بالإضافة إلى نفس تقسيم البيانات بين بيانات للتدريب والاختبار كما في المرحلة السابقة

2-5-2 بناء نموذجي تنبؤ محسن باستخدام خوارزمية LSTM

باستدعاء الدالة المعرفة في ملف (lstm.py) التي تبني النموذج المحسن كما يلي :

```
def build_improved_model(input_dim, output_dim, return_sequences):
    model = Sequential()
    model.add(LSTM(
        input_shape=(None, input_dim),
        units=output_dim,
        return_sequences=return_sequences))

    model.add(Dropout(0.2))

    model.add(LSTM(
        128,
        return_sequences=False))

    model.add(Dropout(0.2))

    model.add(Dense(
        units=1))
    model.add(Activation('linear'))

    return model
```

تعتمد الدالة السابقة على مكتبة keras Long short term memory لتقوم بتطبيق نموذج تنبؤ LSTM وتم زيادة حجم الدفعة Batch_Size من 1 إلى 512 وزيادة العصور Epochs من 1 إلى 20 في النموذج المحسن من الخوارزمية LSTM كما أن الدالة السابقة أضافت إلى عدد العقد في الطبقة المخفية ليصبح 128 بعد أن كان 100 وأضافت معدل انخفاض dropout بمقدار 0.2 لكافة الطبقات. ولبناء النموذج :

```
# Build an improved LSTM model
batch_size = 512
epochs = 20

model = lstm.build_improved_model(X_train.shape[0] - 1, output_dim=unroll_length, return_sequences=True)
```

3-5-2 تدريب النموذج المحسن من LSTM

```
def train_improved_LSTM_model(self):
    from Stock_Price_Predictor import model, X_train, y_train, batch_size, epochs
    model.fit(X_train,
              y_train,
              batch_size=batch_size,
              epochs=epochs,
              verbose=2,
              validation_split=0.05
              )
```

4-5-2 اختبار المتانة Robustness Check

لاختبار متانة النموذج النهائي ومدى فاعلية التنبؤ تم استخدام بيانات جديدة لم يتم تقديمها للنموذج مسبقاً لا من خلال التدريب أو الاختبار وهذه البيانات هي من شركة ألفابيت من الفترة

```
trainScore = model.evaluate(X_train, y_train, verbose=0)

testScore = model.evaluate(X_test, y_test, verbose=0)

range = [np.amin(stocks_data['Close']), np.amax(stocks_data['Close'])]

# Calculate the stock price delta in $

true_delta = testScore * (range[1] - range[0])
data = pd.read_csv('unseenyahosp500.csv') # this is unseen data

stocks = ppd.remove_data(data)

stocks = ppd.get_normalised_data(stocks)

stocks = stocks.drop(['Item'], axis=1)
```

2-5-5 توابع لظهار التنبؤ بسعر الاغلاق

التابع الاول يقوم باظهار قيم سعر الاغلاق مع قيمة التنبؤ به من اللحظة الحالية الى فترة سابقة حوالي الشهر يقوم هذا التابع ب جلب الداتا ويقوم بعملية معالجة عليها بسحب سعر الاغلاق فقط مع ايجاد قيمة التنبؤ

```
data = pd.read_csv('data_preprocessed.csv')
stocks = ppd.remove_data(data)
stocks = ppd.get_normalised_data(stocks)
stocks = stocks[['Close']]
```

```
future_days = 25
stocks['Prediction'] = stocks[['Close']].shift(-future_days)
self.textEdit_10.setText(str(stocks.head(28)))
```

التابع الثاني هو الذي يقوم بعملية التنبؤ لسعر الاغلاق ليوم الغد من اليوم الحالي يقوم ايضا بالفلتر لسحب قيمة سعر الاغلاق فقط وتحويله الى مصفوفة numpy وعملية تطبيع لجعل القيم بين 0-1 من خلال تابع MinMaxScaler ثم يجري عملية تدريب للداتا عن طريق استخدام LSTM, Dense ثم يقوم من خلال سعر الاغلاق لمدة 60 يوم سابق ومن خلال المتحول x_test و التحويل الى مصفوفة numpy والمعالجة يقوم بالتنبؤ بسعر الاغلاق لليوم التالي

```
# Create a new dataframe with only the 'Close' column
data = stocks.filter(['Close'])
# Converting the dataframe to a numpy array
dataset = data.values
# Get /Compute the number of rows to train the model on
training_data_len = math.ceil(len(dataset) * .8)
# Scale the all of the data to be values between 0 and 1
scaler = MinMaxScaler(feature_range=(0, 1))
scaled_data = scaler.fit_transform(dataset)
train_data = scaled_data[0:training_data_len, :]
x_train = []
y_train = []
for i in range(60, len(train_data)):
    x_train.append(train_data[i - 60:i, 0])
    y_train.append(train_data[i, 0])
model = Sequential()

model.add(LSTM(units=50, return_sequences=False))
model.add(Dense(units=25))
model.add(Dense(units=1))
# Create a new dataframe
new_df = stocks.filter(['Close'])
# Get teh last 60 day closing price
last_60_days = new_df[-60:].values
# Create an empty list
X_test = []
# Append teh past 60 days
X_test.append(last_60_days)
# Convert the X_test data set to a numpy array
X_test = np.array(X_test)
# Reshape the data
X_test = np.reshape(X_test, (X_test.shape[0], X_test.shape[1], 1))
# Get the predicted scaled price
pred price = model.predict(X_test)
```

الفصل السادس

التنفيذ والاختبارات

1- التنفيذ

يتم في هذه المرحلة تنويع الجهود المبذولة في المراحل السابقة وذلك بتنفيذ الإجراءات والمهام التي تم تحقيقها آنفاً.

1-1- التقانات المستخدمة

كي نقوم بتنفيذ المشروع وباعتباره مشروع ذكاء صناعي كان لا بد من اختيار التقنيات الملائمة لهذا النوع والتي تحقق كفاءة عالية وقابلية تطبيق ودعم كافي للمجالات المطلوبة وهناك عدة أدوات ساعدت في إنجاز هذه المهمة:

- بيئة البرمجة Pycharm وهي البيئة التي تم تنفيذ المشروع بالكامل ضمنها
- لغة البرمجة Python لما تمتلكه من خصائص ومكتبات داعمة مرتبطة بالذكاء الصناعي بشكل مباشر بالإضافة لسرعتها ومرونتها في تنفيذ الأهداف المطلوبة والعديد من المميزات الأخرى
- مكتبات عديدة من بايثون كانت مفصلية في إنجاز المشروع وأهمها Numpy, Pandas, Keras, TensorFlow, Sklearn على اختلاف مهماتها
- أداة Jupyter Notebook كانت الوسيلة الأولى في إنجاز المهام الأساسية المطلوبة من المشروع قبل الانتقال إلى تصميم الواجهات النهائية وربطها بالمشروع
- مكتبة PYQT5 التي استخدمت في تصميم الواجهات

2- إنجاز الإجراءات

هدف هذا المشروع هو التنبؤ الدقيق بأسعار الإغلاق المستقبلية لأسهم شركات معينة على فترات زمنية مستقبلية ولتحقيق هذه المشروع تم استخدام خوارزمية تحت اسم LSTM للتنبؤ بأسعار أسهم شركة ألفا وأسهم S&P 500 باستخدام بيانات تاريخية وفيما يلي توضيح التنفيذ للإجراءات والنتائج عن كل مرحلة:

	A	B	C	D	E	F	G
	Date	High	Low	Open	Close	Volume	Adj Close
	2000/01/03	1478	1438.36	1469.25	1455.22	9.32E+08	1455.22
	2000/01/04	1455.22	1397.43	1455.22	1399.42	1.01E+09	1399.42
	2000/01/05	1413.27	1377.68	1399.42	1402.11	1.09E+09	1402.11
	2000/01/06	1411.9	1392.1	1402.11	1403.45	1.09E+09	1403.45
	2000/01/07	1441.47	1400.73	1403.45	1441.47	1.23E+09	1441.47
	2000/01/10	1464.36	1441.47	1441.47	1457.6	1.06E+09	1457.6
	2000/01/11	1458.66	1434.42	1457.6	1438.56	1.01E+09	1438.56
	2000/01/12	1442.6	1427.08	1438.56	1432.25	9.75E+08	1432.25
0	2000/01/13	1454.2	1432.25	1432.25	1449.68	1.03E+09	1449.68
1	2000/01/14	1473	1449.68	1449.68	1465.15	1.09E+09	1465.15

بعد حفظ البيانات محلياً في ملف 'yahoosps500.csv' ننتقل إلى الخطوة التالية

1-2-2 معالجة أولية للبيانات بالاستفادة من مكتبة Pandas

حساب المتوسط mean والانحراف المعياري standard deviation والقيمة العظمى Max والدنيا Min والنتائج:

	Date	Open	High	Low	Close	Volume
0	2020/07/17	3224.209961	3233.520020	3205.649902	3224.729980	3993830000
1	2020/07/16	3208.360107	3220.389893	3198.590088	3215.570068	3961230000
2	2020/07/15	3225.979980	3238.280029	3200.760010	3226.560059	4669760000
3	2020/07/14	3141.110107	3200.949951	3127.659912	3197.520020	4476170000
4	2020/07/13	3205.080078	3235.320068	3149.429932	3155.219971	4890780000

Open	---	mean :	1612.8617081157508	Std:	623.7806557037948	Max:	3380.4499509999996	Min:	679.2800292999999
High	---	mean :	1622.2491352942918	Std:	625.4272480150993	Max:	3393.5200200000004	Min:	695.2700195
Low	---	mean :	1602.6339384709174	Std:	621.8154703306687	Max:	3378.830078	Min:	666.789978
Close	---	mean :	1613.0281263259676	Std:	623.840630715595	Max:	3386.149902	Min:	676.5300292999999
Volume	---	mean :	3152598171.4396286	Std:	1520150153.7959082	Max:	11456230000	Min:	356070000

الجدول (17) البيانات بعد حساب الانحراف المعياري

2-2-2 إزالة البيانات غير الضرورية

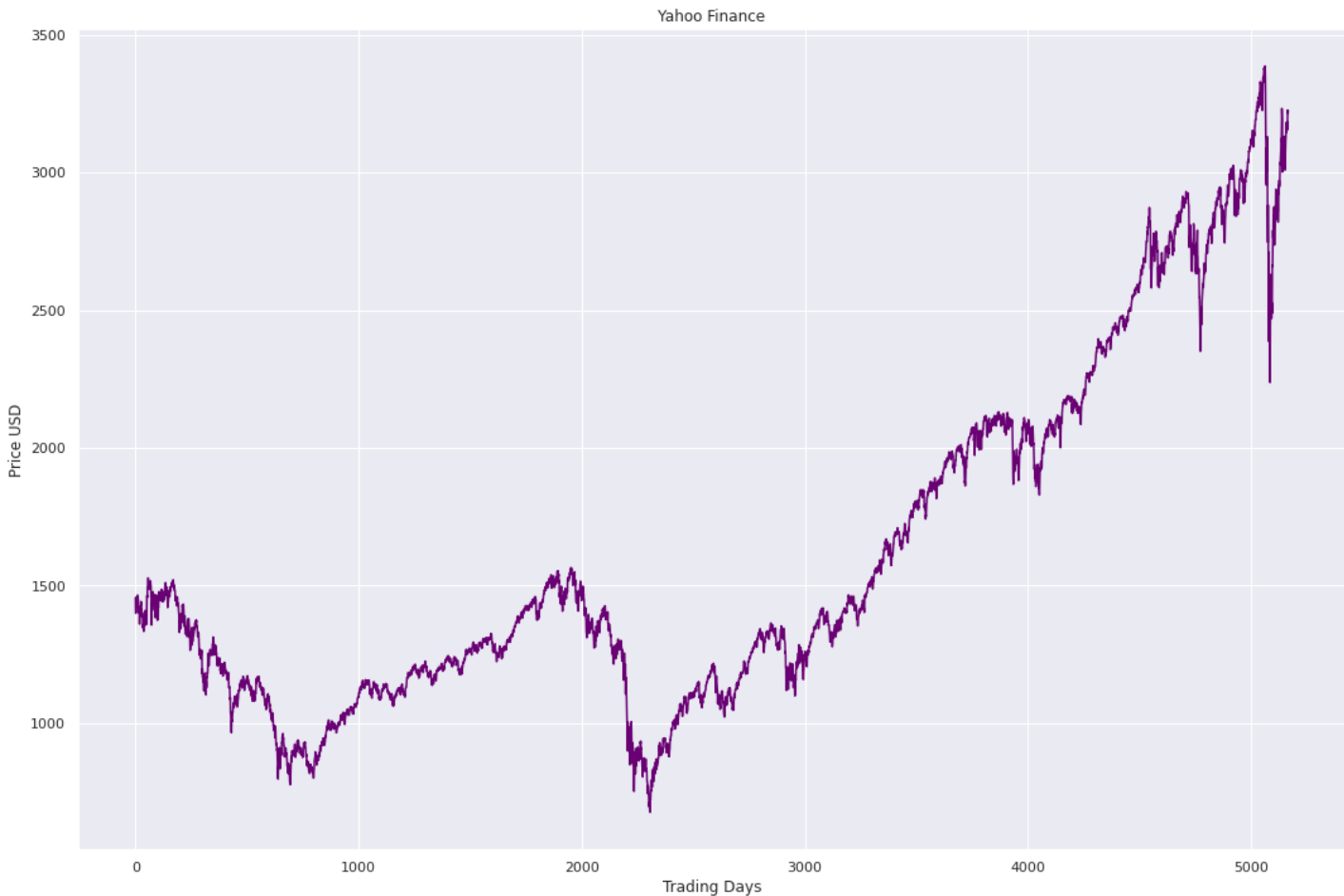
مثل التاريخ وأعظم قيمة وأدنى تحققت في اليوم ويصبح شكل البيانات:

	Item	Open	Close	Volume
0	0	1469.250000	1455.219971	931800000
1	1	1455.219971	1399.420044	1009000000
2	2	1399.420044	1402.109985	1085500000
3	3	1402.109985	1403.449951	1092300000
4	4	1403.449951	1441.469971	1225200000

	Item	Open	Close	Volume
5163	5163	3205.080078	3155.219971	4890780000
5164	5164	3141.110107	3197.520020	4476170000
5165	5165	3225.979980	3226.560059	4669760000
5166	5166	3208.360107	3215.570068	3961230000
5167	5167	3224.209961	3224.729980	3993830000

الجدول (18) يمثل شكل البيانات بعد إزالة المعلومات غير الضرورية منها

3-2-2 تمثيل البيانات الخام رسوماً كما هي في وضعها الحالي



الشكل (20) تمثيل البيانات الأولية

نلاحظ عدد العينات في المحور x أكثر من 5000 والسعر بالدولار الأمريكي في المحور y

4-2-2 تطبيع البيانات باستخدام دالة MinMaxScaler

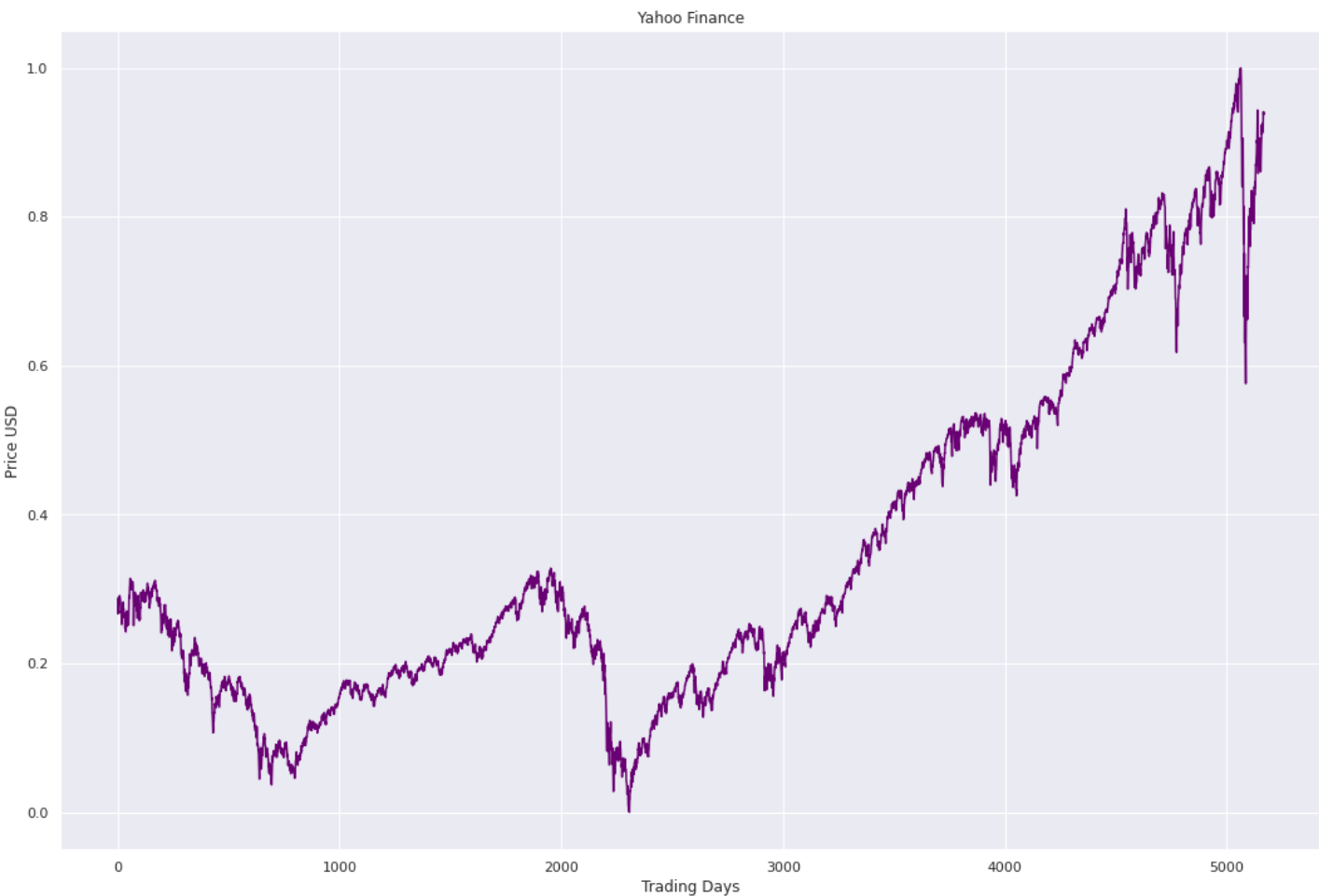
لتحويلها إلى قيم بين الصفر والواحد وحفظ البيانات في ملف جديد yahoo_preprocessed.csv وبإعادة حساب المتوسط mean والانحراف المعياري standard deviation والقيمة العظمى Max والدنيا Min نجد:

	Item	Open	Close	Volume
0	0	0.292455	0.287380	0.051867
1	1	0.287261	0.266787	0.058822
2	2	0.266603	0.267779	0.065713
3	3	0.267599	0.268274	0.066326
4	4	0.268095	0.282305	0.078299

Open	---	mean :	0.3456212329760413	Std:	0.23092980959569337	Max:	1.0	Min:	0.0
Close	---	mean :	0.3456197330339161	Std:	0.23023178896823235	Max:	1.0	Min:	0.0
Volume	---	mean :	0.25193584339681846	Std:	0.13694849027364547	Max:	1.0	Min:	0.0

الجدول (19) يمثل شكل البيانات بعد التطبيع

5-2-2 تمثيل البيانات بعد التطبيع رسوياً



الشكل (21) تمثيل البيانات بعد تطبيق المعالجة الأولية عليها

نلاحظ عدد العينات في المحور x أكثر من 5000 وفي المحور y يتراوح السعر بين 0 و 1

2-3- بناء النموذج المعياري باستخدام التراجع الخطي

سوف يتم في هذه المرحلة بناء النموذج الأولي بالاعتماد على التراجع الخطي والبيانات التي سوف يتم تطبيق النموذج عليها هي البيانات بعد المعالجة الأولية من المرحلة السابقة وهي ممثلة بالشكل (20)

	Item	Open	Close	Volume
0	0	0.292455	0.287380	0.051867
1	1	0.287261	0.266787	0.058822
2	2	0.266603	0.267779	0.065713
3	3	0.267599	0.268274	0.066326
4	4	0.268095	0.282305	0.078299

الجدول (20) البيانات بعد المعالجة الأولية

1-3-2 تقسيم البيانات

يتم تقسيم البيانات إلى بيانات للتدريب وبيانات للاختبار بهدف تطبيق نموذج التراجع الخطي باستدعاء الدالة (المعرفة في ملف stock_data.py) والتي تقوم بتقسيم البيانات من أجل نموذج التراجع الخطي وفيما يلي نتيجة تقسيم البيانات :

```
x_train (3541, 1)
y_train (3541, 1)
x_test (1627, 1)
y_test (1627, 1)
```

2-3-2 تدريب النموذج الأولي

في هذه المرحلة يتم بناء النموذج بالاعتماد على مكتبة scikit-learn linear_model واستدعاء الدالة المعرفة في ملف LinearRegressionModel.py والتي تقوم ببناء النموذج الخاص بالمشروع. [4]

3-3-2 الحصول على التنبؤ لمجموعة بيانات الاختبار

4-3-2 التمثيل الرسومي لنتائج التنبؤ باستخدام النموذج الأولي باستخدام التراجع الخطي مقارنة بالقيم الفعلية



الشكل (22) التراجع الخطي ويظهر التنبؤ باللون الأزرق والقيم الفعلية باللون الأخضر

X تمثل أيام التداول – Y تمثل سعر سهم الإغلاق بالدولار الأمريكي

الخط الأخضر يمثل سعر الإغلاق – الخط الأزرق يمثل سعر الإغلاق المتوقع به

Train Score: 0.1852 MSE (0.4303 RMSE)

Test Score: 0.08133781 MSE (0.28519784 RMSE)

5-3-2 حساب دقة التنبؤ ونتيجة الاختبار لنتائج النموذج الأولي

لحساب نتيجة التدريب والاختبار نقوم بحساب متوسط مربع الخطأ (MSE) Mean Squared Error ومن ثم جذر

متوسط مربع الخطأ Root Mean Squared Error

والنتائج:

Train Score: 0.2114 MSE (0.4598 RMSE)

Test Score: 0.08496722 MSE (0.29149136 RMSE)

4-2- بناء نموذج التنبؤ باستخدام خوارزمية LSTM

في هذه الخطوة سوف يتم بناء نموذج التنبؤ بالاعتماد على خوارزمية تعلم الآلة تحت اسم الذاكرة طويلة قصيرة المدى Long-Sort Term Memory اختصاراً (LSTM) وفي هذا الجزء سوف يتم استخدام الخوارزمية للتدريب والاختبار على مجموعة البيانات.

1-4-2 بناء نموذج تنبؤ أولي باستخدام LSTM

بغرض بناء النموذج الأولي لخوارزمية LSTM سوف نقوم بتضمين مكتبة Keras لبناء النموذج وبلاستفادة من مكتبة TensorFlow والبيانات التي سوف يتم العمل عليها ذاتها البيانات بعد المعالجة الأولية وهي ممثلة بالجدول (20) الموضوع سابقاً

2-4-2 تقسيم البيانات إلى بيانات للتدريب وبيانات للاختبار واختبار البيانات لنموذج lstm

وبالتالي يتم تقسيم العينات (البيانات) كما يلي

```
x_train (4612, 50, 3)
y_train (4612,)
x_test (446, 50, 3)
y_test (446,)
```

3-4-2 بناء نموذج تنبؤ أولي باستخدام الذاكرة طويلة قصيرة المدى

ونتيجة حساب وقت التنفيذ

```
compilation time : 0.01821136474609375
```

4-4-2 تدريب النموذج الأولي من خوارزمية LSTM

بالاعتماد على مكتبة TensorFlow سوف يجري تدريب النموذج عملية التدريب:

```
4381/4381 [=====] - 172s 39ms/step - loss: 0.0012 - val_loss: 0.0032
```

5-4-2 اجراء عملية التنبؤ باستخدام بيانات الاختبار

6-4-2 تمثيل نتيجة التنبؤ باستخدام النموذج الأولي رسومياً



الشكل (23) التمثيل البياني للنموذج الأولي لخوارزمية LSTM



الشكل (24) التمثيل البياني للنموذج المحسن لخوارزمية LSTM

7-4-2 حساب دقة التنبؤ ونتيجة الاختبار لنتائج النموذج الأولي من خوارزمية LSTM

لحساب نتيجة التدريب نقوم بحساب متوسط مربع الخطأ Mean Squared Error (MSE) ومن ثم جذر متوسط مربع الخطأ Root Mean Squared Error :

```
Train Score: 0.00031041 MSE (0.01761858 RMSE)
Test Score: 0.00199818 MSE (0.04470103 RMSE)
```

5-2- بناء النموذج المحسن من خوارزمية LSTM

تهدف هذه الخطوة إلى بناء نموذج تنبؤ بالاعتماد على ذات الخوارزمية ولكن بهدف تحقيق نتائج أفضل وفق الخطوات التالية [5] :

1-5-2 تقسيم البيانات إلى بيانات للتدريب وبيانات للاختبار واختبار البيانات لنموذج lstm

ونلاحظ أنه تم استخدام نفس البيانات بعد المعالجة الأولية والتي تم استخدامها في بناء النموذجين السابقين بالإضافة إلى نفس تقسيم البيانات بين بيانات للتدريب والاختبار كما في المرحلة السابقة

```
x_train (4612, 50, 3)
y_train (4612,)
x_test (446, 50, 3)
y_test (446,)
```

2-5-2 بناء نموذجي تنبؤ محسن باستخدام خوارزمية LSTM

تعتمد الدالة على مكتبة keras Long short term memory لتقوم بتطبيق نموذج تنبؤ LSTM وتم زيادة حجم الدفعة Batch_Size من 1 إلى 512 وزيادة العصور Epochs من 1 إلى 20 في النموذج المحسن من الخوارزمية LSTM

كما أن الدالة السابقة أضافت إلى عدد العقد في الطبقة المخفية ليصبح 128 بعد أن كان 100 وأضافت معدل انخفاض dropout بمقدار 0.2 لكافة الطبقات.
وقت البناء:

```
compilation time : 0.020635128021240234
```

3-5-2 تدريب النموذج المحسن من LSTM

ويستغرق تدريب النموذج وقتاً بطبيعة الحال:

```

9/9 - 11s - loss: 4.3423e-04 - val_loss: 0.0010
Epoch 10/20
9/9 - 11s - loss: 4.1352e-04 - val_loss: 0.0010
Epoch 11/20
9/9 - 11s - loss: 4.2082e-04 - val_loss: 9.8191e-04
Epoch 12/20
9/9 - 12s - loss: 4.2834e-04 - val_loss: 0.0010
Epoch 13/20
9/9 - 12s - loss: 4.0217e-04 - val_loss: 9.8623e-04
Epoch 14/20
9/9 - 12s - loss: 4.1540e-04 - val_loss: 8.8144e-04
Epoch 15/20
9/9 - 12s - loss: 4.0718e-04 - val_loss: 8.1230e-04
Epoch 16/20
9/9 - 12s - loss: 3.8957e-04 - val_loss: 9.0547e-04
Epoch 17/20
9/9 - 12s - loss: 3.8967e-04 - val_loss: 9.7957e-04
Epoch 18/20
9/9 - 12s - loss: 4.0217e-04 - val_loss: 8.4695e-04
Epoch 19/20
9/9 - 12s - loss: 3.9935e-04 - val_loss: 7.8799e-04
Epoch 20/20
9/9 - 12s - loss: 3.9314e-04 - val_loss: 7.6090e-04

```

4-5-2 حساب دقة التنبؤ ونتيجة الاختبار للنموذج المحسن من خوارزمية LSTM

لحساب نتيجة التدريب نقوم بحساب متوسط مربع الخطأ (MSE) Mean Squared Error ومن ثم جذر متوسط

مربع الخطأ Root Mean Squared Error

Test Score	0.003010204993188381
RMSE	0.05486533507770076

Train Score: 0.00024132 MSE (0.01553438 RMSE)
 Test Score: 0.00275816 MSE (0.05251817 RMSE)

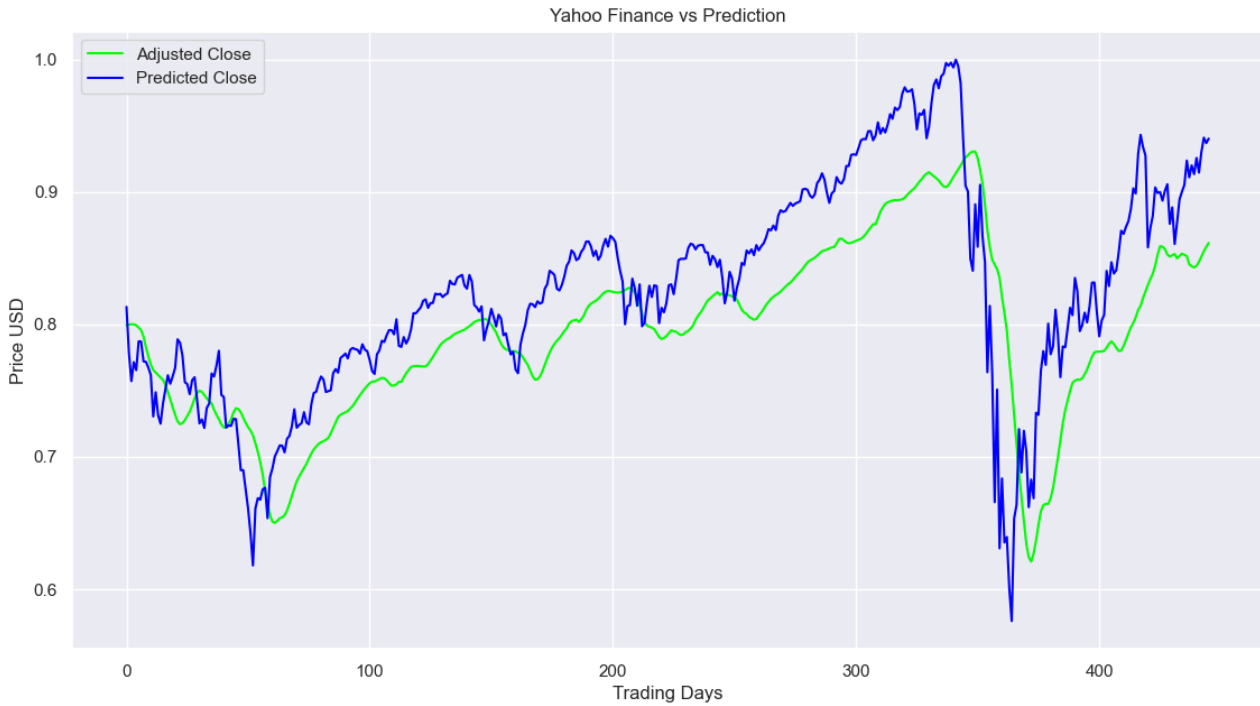
الشكل (26) نتائج تقييم النموذج المحسن

3- التحسين

لبناء النموذج المحسن من خوارزمية LSTM تم العمل على تحسين المعايير للخوارزمية Parameters للحصول على نتائج تنبؤ أفضل. وتم تحقيق التحسين باختبار وتحليل كل معيار على حدى واختيار القيمة النهائية لكل منهم ولتحسين النموذج تم اتباع الخطوات التالية:

- زيادة عدد العقد المخفية في كل طبقة من 100 إلى 128
- تم إضافة Dropout بمقدار 0.2 في كل طبقة من طبقات LSTM
- زيادة حجم الدفعة Batch_Size من 1 إلى 512
- زيادة العصور Epochs من 1 إلى 20
- إضافة مفصل verbose = 2
- إنجاز التنبؤ بحجم الدفعة

وهذه الخطوات أدت إلى تحسن مباشر في MSE لمجموعات الاختبار ويمكن ملاحظة الفرق بالتمثيل الرسومي للخوارزمية المحسنة



الشكل (27) تمثيل الخوارزمية المحسنة LSTM

4- النتائج

4-1- تطوير النموذج والتحقق من صلاحيته

مع كل نموذج تم تطويره من النماذج السابقة تم تحسين نتائج التنبؤ والتقليل من متوسط مربع الخطأ Mean Squared Error بشكل ملاحظ.

4-2- اختبار المتانة Robustness Check

لاختبار متانة النموذج النهائي ومدى فاعلية التنبؤ تم استخدام بيانات جديدة لم يتم تقديمها للنموذج مسبقاً لا من خلال التدريب أو الاختبار وهذه البيانات هي من شركة ألفابيت

	Open	Close	Volume
0	0.047007	0.266938	0.062420
1	0.411130	0.307189	0.534262
2	0.298675	0.445190	0.294750
3	0.436814	0.147659	0.323604
4	0.000000	0.000000	0.000000
5	0.010373	0.175304	0.056456
6	0.129167	0.020715	0.108798
7	0.071128	0.315592	0.639255
8	0.108504	0.225507	0.293784
9	0.427018	0.409069	1.000000
10	0.573639	0.582235	0.612536
11	0.587304	0.669960	0.594563
12	0.813288	0.826686	0.685003
13	0.861035	0.984371	0.304812
14	1.000000	1.000000	0.171766
(14, 1, 3)			
(14,)			
Test Score: 0.2547 MSE (0.5047 RMSE)			

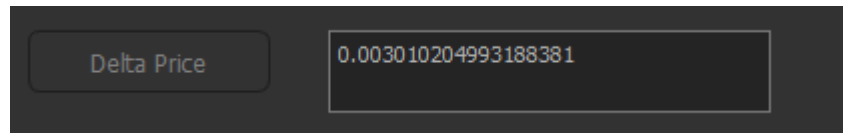
الجدول (21) نتيجة البيانات المستخدمة لاختبار متانة النموذج

4-3- التبرير

بمقارنة مع النموذج المبدئي الذي استخدم التراجع الخطي في التنبؤ مع النموذج النهائي المحسن من خوارزمية LSTM تم ملاحظة تغير في MSE متوسط مربع الخطأ وهذا الانخفاض الواضح في معدل الخطأ يوضح أن النموذج النهائي قد تجاوز النموذج المبدئي والأساسي في الأداء.

بالإضافة إلى أن السعر الوسطي دلتا Delta بين البيانات الفعلية والبيانات التي تم التنبؤ بها كانت

Delta Price: 0.000931 - RMSE * Adjusted Close Range



الشكل (28) سعر دلتا الوسطي

وهو ما يقارب أقل من سنت واحد من الاختلاف

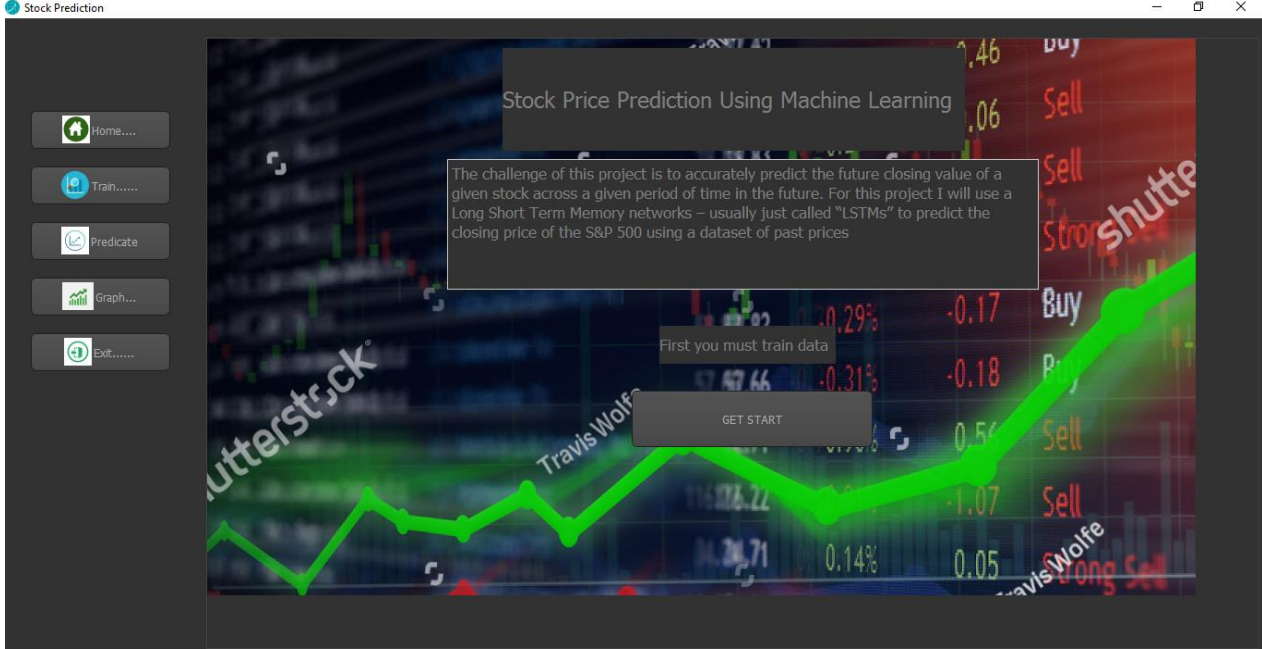
الفصل السابع

دليل الاستخدام

1- استعراض الواجهات الرئيسية

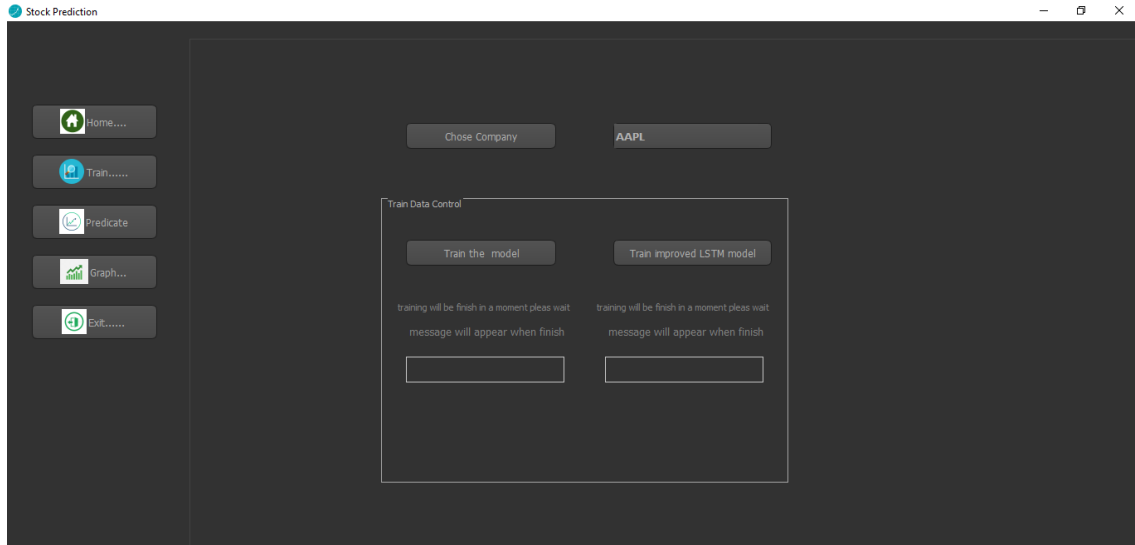
• النافذة الرئيسية Home

تعرض للزائر تعريف بالبرنامج وهدفه



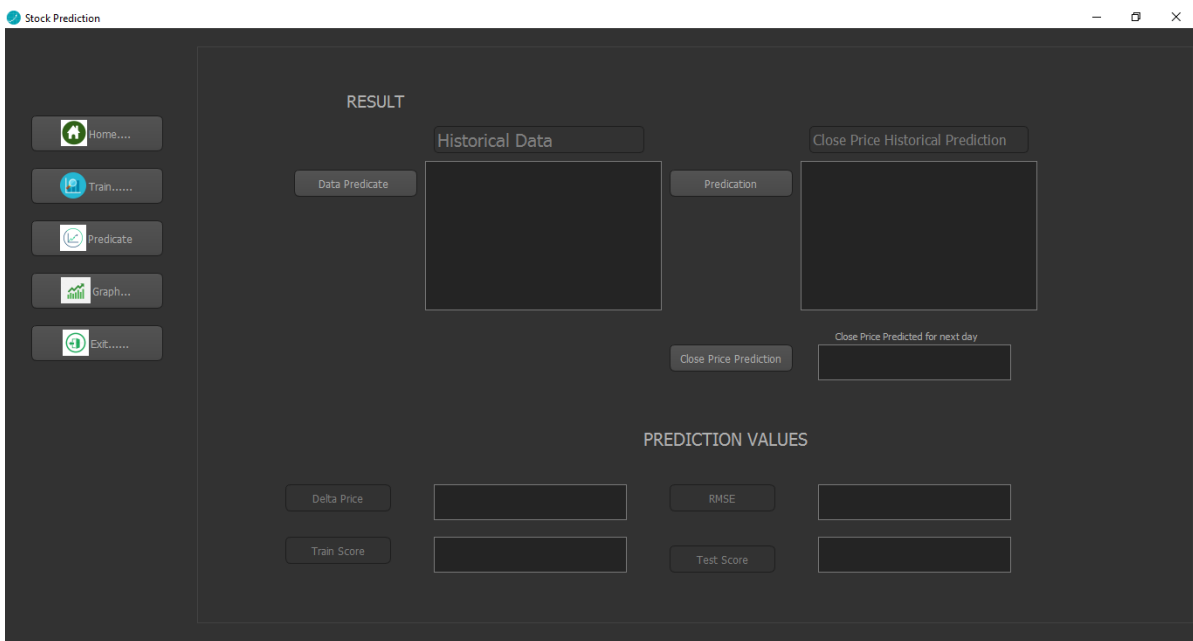
الشكل (29) الواجهة الرئيسية للبرنامج

- النافذة الثانية هي الخاصة بإجراء عملية اختيار شركة التداول واستحصال البيانات وتدريب الداتا وتظهر كما يلي للمستخدم عند تشغيل البرنامج
- يتم اختيار الشركة من شريط combobox كما يمكن اضافة اسم شركة غير موجودة وعند الضغط على زر chose company يتم استحصال البيانات من yahoo finance وحفظها بملف data_preprocessed
- عند ضغط المستخدم على زر تدريب النموذج Train The model يجري تدريب نموذج التراجع الخطي
- وعند الضغط على زر تدريب نموذج الخوارزمية Train Improved LSTM model يجري تدريب النموذج المحسن من الخوارزمية
- عند الانتهاء من عملية التدريب تظهر رسالة اشعار بالانتهاء



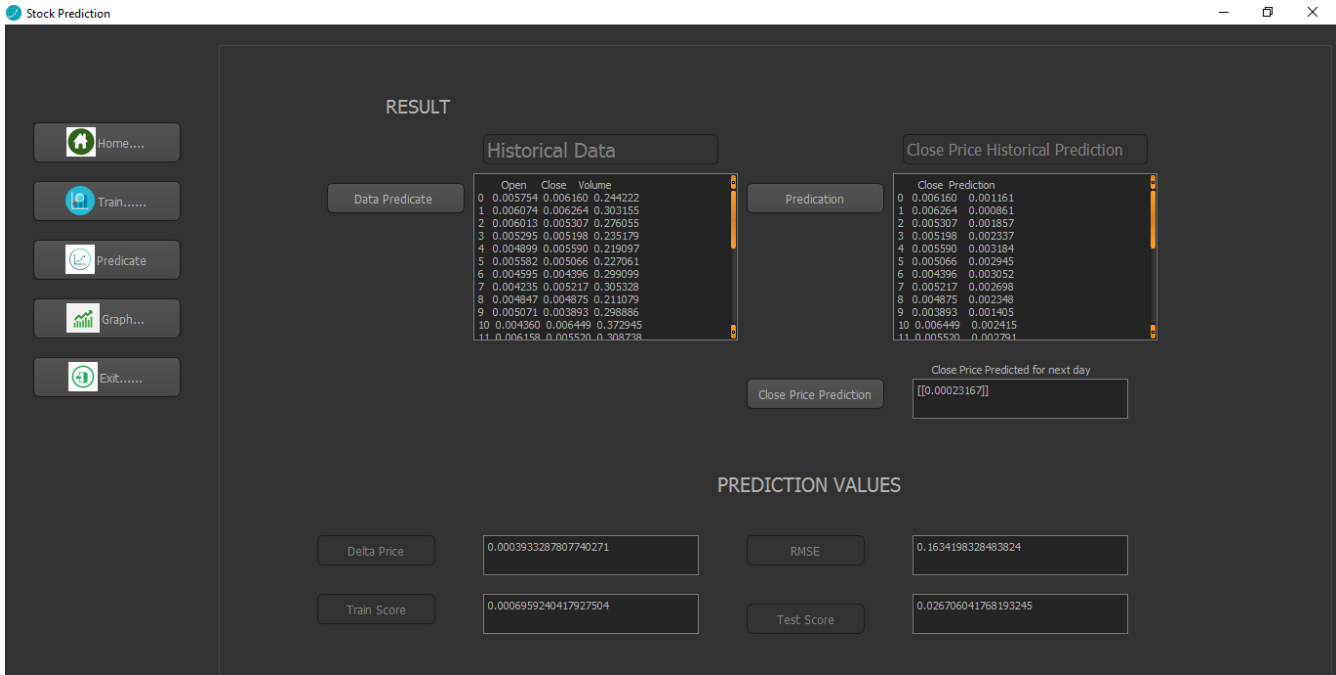
الشكل (30) واجهة التدريب

- في النافذة الثالثة يتم في مربع النص Historical Data عرض الداتا المستخدمة في التنبؤ عند الضغط على الزر predicate وايضا عرض قيم Delta price, Train score, Test Score, RMSE,
- في ربع نص Close Price Historical يتم عرض قيم سعر الاغلاق من اللحظة الحالية الى ما قبل شهر مع عرض قيم التنبؤ لسعر الاغلاق لتلك الايام السابقة
- وفي مربع نص Close Price Predicted next day يتم عرض التنبؤ لسعر الاغلاق لليوم الغد من اليوم الحالي



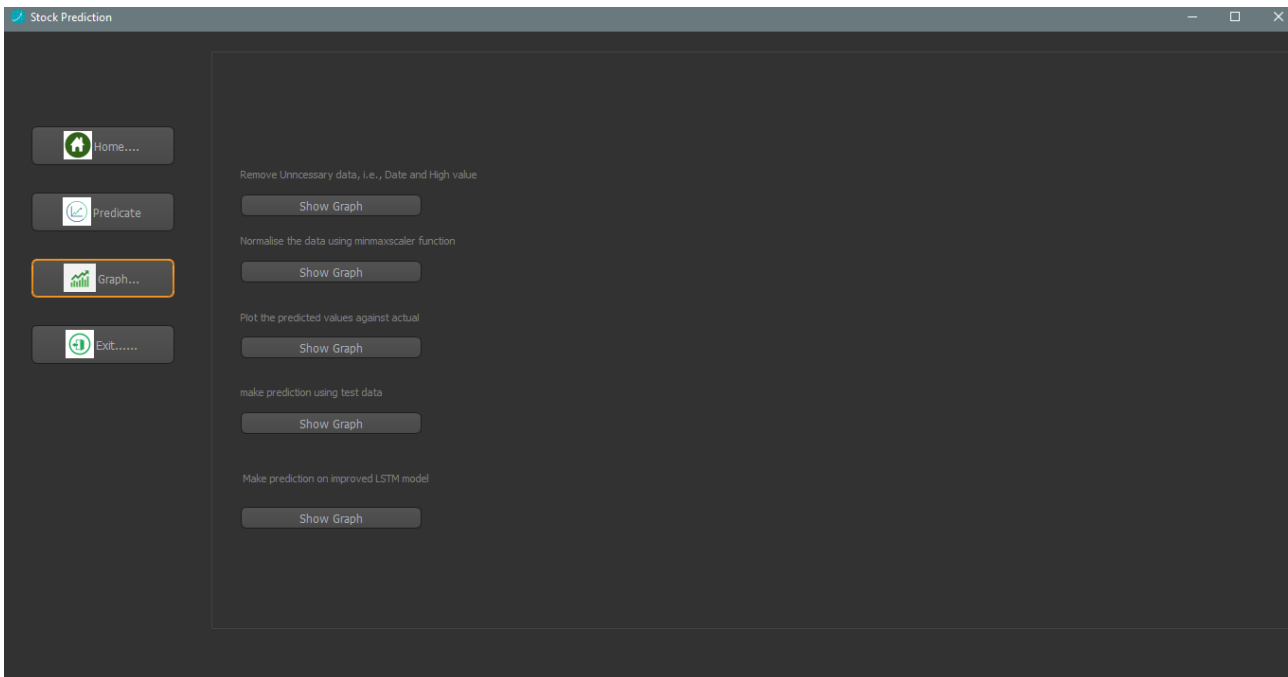
الشكل (31) واجهة التنبؤ

صورة للخروج



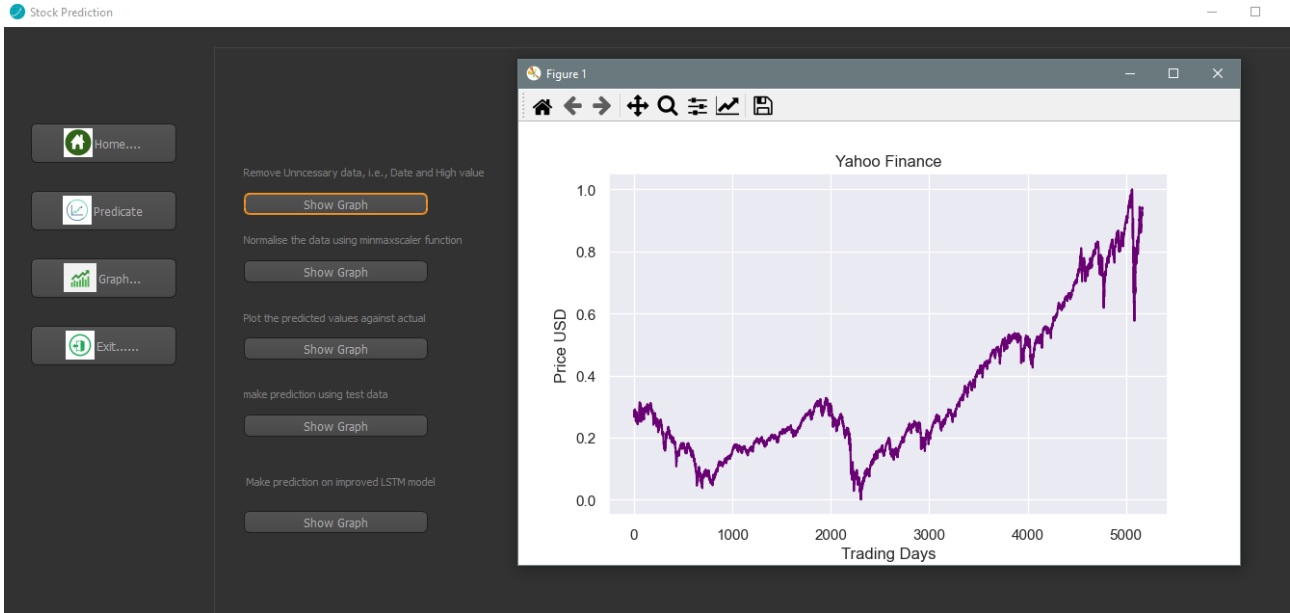
الشكل (32) واجهة التنبؤ بعد تنفيذ التنبؤ وإظهار القيم

- في النافذة الثالثة تحت اسم Graph يوجد عدة خيارات لعروض التمثيل الرسومي للبيانات البدائية وعند إنجاز تدريب النماذج يمكن عرض نتائج التنبؤ رسوياً



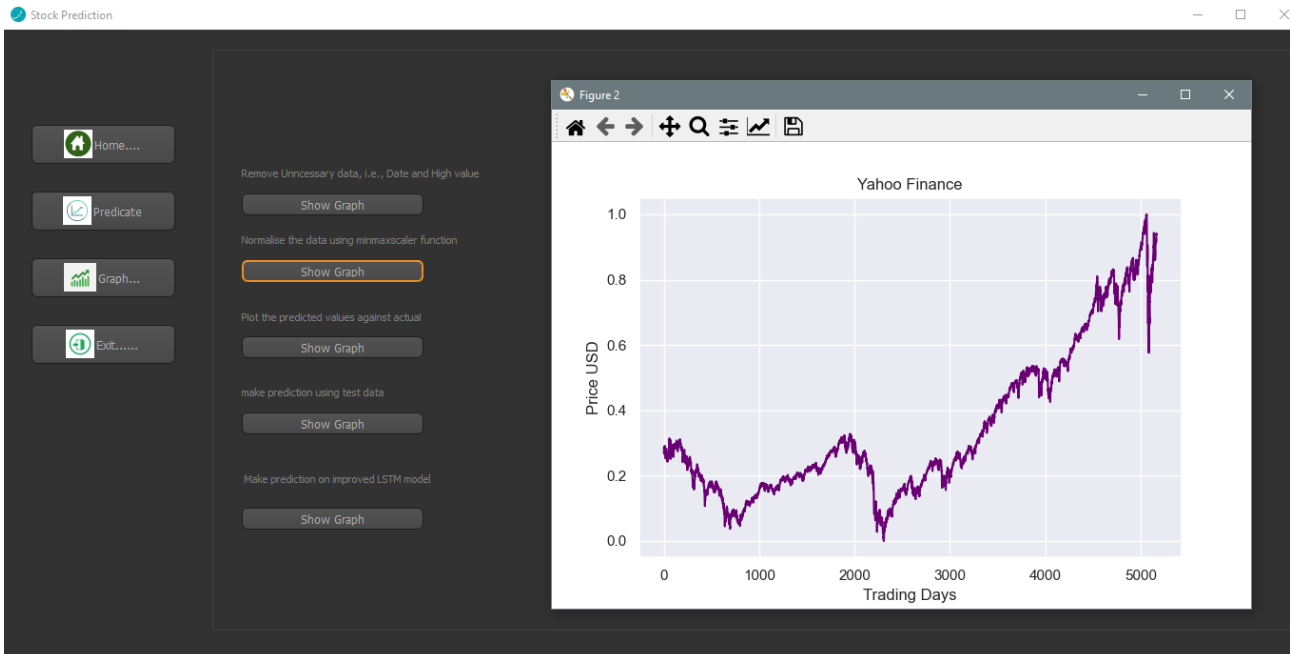
الشكل (33) واجهة إظهار المخططات

عند الضغط على الزر الأول يتم عرض البيانات كاملة



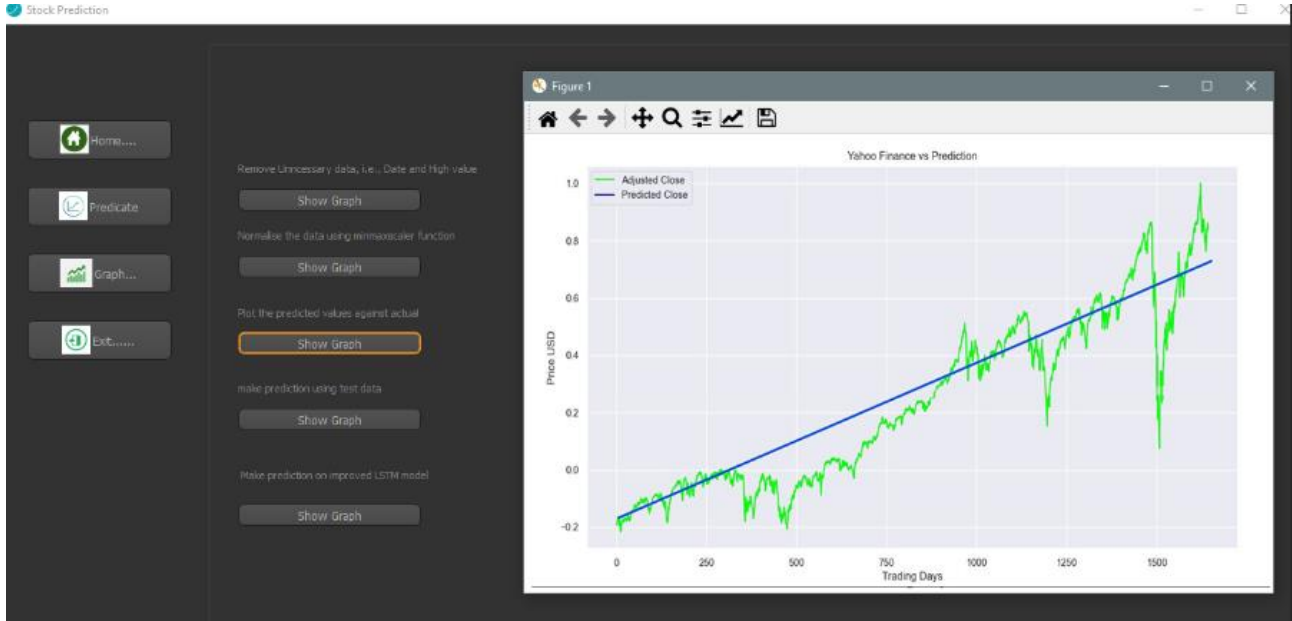
الشكل (34) واجهة إظهار التمثيل الرسومي للبيانات بعد المعالجة الأولية

عند الضغط على الزر الثاني يتم عرض البيانات بعد اجراء المعالجة الأولية عليها



الشكل (35) البيانات بعد اجراء المعالجة الأولية عليها

عند الضغط على الزر الثالث يتم عرض نتائج التنبؤ للنموذج المعياري وهو نموذج التراجع الخطي



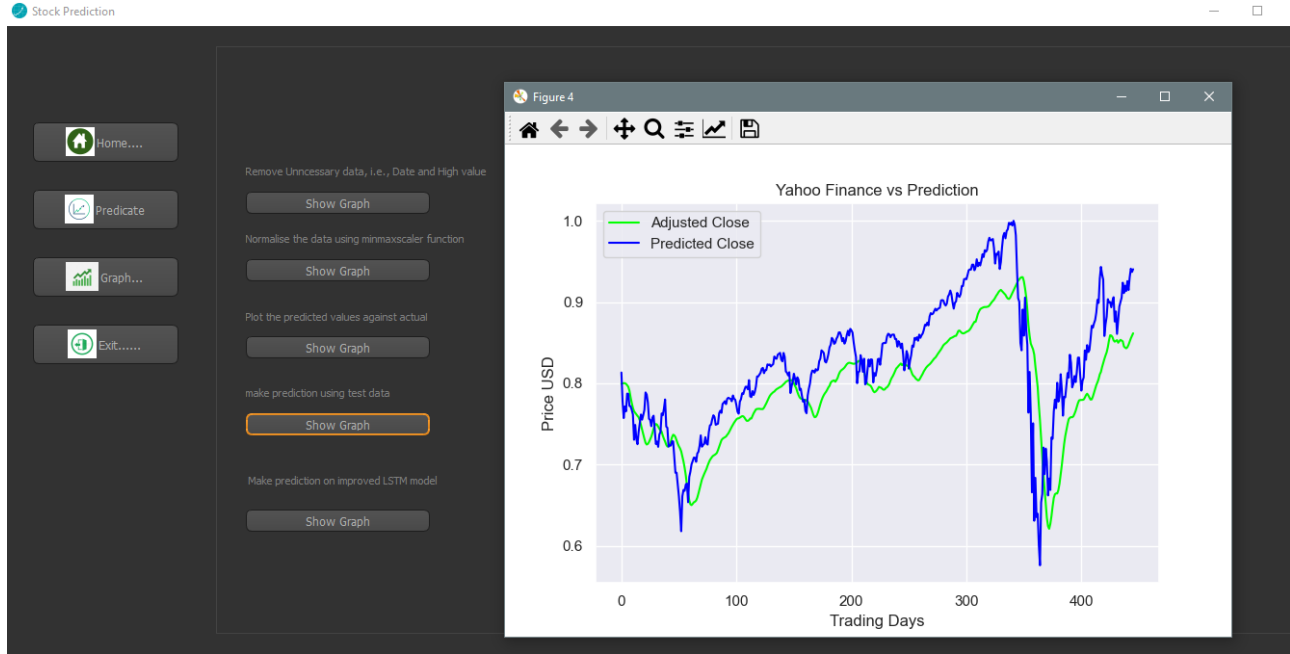
الشكل (36) التراجع الخطي

عند الضغط على الزر الرابع يتم عرض نتائج التنبؤ للنموذج المعتمد على خوارزمية LSTM الأولية



الشكل (37) خوارزمية التنبؤ LSTM الأولية

عند الضغط على الزر الرابع يتم عرض نتائج التنبؤ للنموذج المعتمد على خوارزمية LSTM المحسنة



الشكل (38) خوارزمية التنبؤ LSTM المحسنة

الخاتمة والآفاق المستقبلية

1- الخاتمة

أفضل النهايات دوماً تكون بدايات لما يليها وهذا ما يمكن قوله عن هذا الجهد المبذول، لقد بدأنا هذا العمل بهدف البحث في مجالات استخدام تقنيات تعلم الآلة وكان الغرض محدداً وهو إمكانية تحقيق تنبؤ دقيق لقضية مستقبلية، وتحقيق الاستغلال الأمثل للبيانات الكبيرة المتوافرة والخورازميات والطرق المتعددة لتحقيق أهداف يصعب حصرها لكثرتها، وانتهينا كما يهدف كل مصمم نظام أو مطور برمجيات في أي عمل يؤديه أن يسعى لحل مشكلة موجودة بأنسب الطرق.

لقد تطور قطاع التكنولوجيا اليوم إلى مرحلة أصبح فيها وجوده ضرورياً وحتمياً في كافة المجالات المتعلقة بحياة الإنسان، وأضحت تقنيات الذكاء الصناعي لا غنى عنها في تحسين الأداء والانطلاق في مرحلة جديدة من تطوير البرمجيات التي تتخطى المفهوم التقليدي لحل المشكلات وتطرح طرقاً غير معتادة لتحقيق أهداف لم يكن من المتوقع إنجازها بسهولة، مع الأخذ بعين الاعتبار بأن هذا العمل كان مبنياً بالضرورة على جهد الكثيرين من العاملين في هذا المجال ممن ساهموا في تطوير هذه التقنيات وإتاحتها للباحثين وتطويرها وصيانتها باستمرار ونأمل أن يكون هذا العمل خطوة واحدة للأمام في طريق ما زال يبدو طويلاً ولكنه يدعو للتفاؤل بما يكفي لاستمرار العمل فيه.

إن النتائج الإيجابية التي يقدمها هذا المشروع في مجال التنبؤ بأسعار الأسهم المستقبلية كنتيجة نهائية والتحسين المستمر في كل مرحلة كان محركاً للفضول ودافعاً للاستكشاف أكثر والاختبار أكثر على الإمكانيات الواسعة التي يقدمها هذا المشروع والمجال الواسع للتطوير والإبداع فيه.

2- الآفاق المستقبلية

مع أن المشروع كان يهدف بالدرجة الأولى للتغلب على المستقبل وإخضاعه، إلا أنه لا بد من التواضع أمامه كون ما لا نعرفه سوف يبقى أكثر مما نعرفه ولن يتغير هذا قريباً، وبالتالي فإن المستقبل مفتوح على مصراعيه على إمكانيات التطوير الكثيرة التي يمكن العمل عليها ومنها مثلاً:

- 1- توفير مجال أوسع من الخيارات ليتم التنبؤ بها بأسعار الأسهم عوضاً عن التركيز على شركة واحدة
- 2- توفير إمكانية التنبؤ لمجال زمني مستقبلي أوسع من أسابيع إلى أشهر وحتى سنوات مع تقديم النصائح المباشرة للمستثمر بالشراء أو البيع
- 3- الأخذ بعين الاعتبار الأخبار اليومية والمعلومات النصية والاستفادة من تقنيات تعلم آلي أخرى يمكن تطبيقها في معالجة النصوص وإضافة طبقة جديدة هامة تزيد من دقة التنبؤ والأخذ بعين الاعتبار التقلبات المفاجئة التي قد يكون سببها إما تنظيمياً داخل الشركة أو سياسياً أو اقتصادياً وغيرها
- 4- تطوير نماذج تنبؤ إضافية حتى يمكن تحقيق دقة أقرب ما يمكن إلى المثالية ولمجال
- 5- توفير واجهة مستخدم أكثر تفاعلية وحيوية وسهولة في الاستخدام ويمكن أيضاً توفير تطبيق للهواتف

المراجع

Python Machine Learning by Sebastian Raschka 2nd Edition [1]

Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and [2]

Techniques to Build Intelligent Systems 1st Edition

[/http://colah.github.io/posts/2015-08-Understanding-LSTMs](http://colah.github.io/posts/2015-08-Understanding-LSTMs) [3]

<https://realpython.com/linear-regression-in-python> [4]

[/https://keras.io/api/layers/recurrent_layers](https://keras.io/api/layers/recurrent_layers) [5]

<https://bradleyboehmke.github.io/HOML/intro.html> [6]

[/https://stackoverflow.com](https://stackoverflow.com) [7]