المجلة الدولية للهندسة وتقنية المعلومات
journal homepage:www.ijeit.misuratau.edu.ly

التنبؤ بأسعار الذهب في السوق المحلية باستخدام نموذج دعم الاتجاهات SVM استنادا على تحليل الآراء

عبد الباسط مفتاح الاجنف
الاكاديمية الليبية
basitegnaf@Gmail.com

محمد احمد الصلابي
جامعة مصراته كلية تقنية المعلومات
m.sullabi@it.misuratau.edu.ly

الملخص-- **لقد كان التنبؤ بأسعار الذهب دائمًا موضوعًا بحثيًا مهمًا وصعبًا بالنسبة لمستثمري الذهب. في النماذج التقليدية، استخدم معظم العلماء سعر الذهب التاريخي أو المؤشرات الاقتصادية للتنبؤ بأسعار الذهب. ومن أجل حل معظم مشاكل النماذج التقليدية، يجب التنقيب في أخبار الذهب على شبكة الإنترنت والتي قد تكون موضوعاً لاستخدام أسلوب تحليل الآراء، الذي يهدف إلى التأكد من الاستقطاب في الرأي والكلمات في أخبار الذهب اليومية. تمثل درجة الرأي اتجاهات السوق الحالية، لذلك تستخدم هذه الدراسة درجة الرأي من خلال التنقيب عن النص كمؤشر مدخل للنموذج المقترح في التنبؤ بأسعار الذهب مع اسعار العملات النقدية كمؤشرات اقتصادية. تم الاعتماد في بحثنا على قاعدة بيانات تحتوي على أكثر من 102000 تغريده من موقع X (تويتر)، تم تحميلها على فترات متفاوتة لمدة عام تقريبا من سنة 2021 الى 2022. تم تدريب قاعدة البيانات بناء على مفهوم تحليل الآراء باستخدام قواعد المعجم. والتي تقوم بمعالجة المسبقة للبيانات ومقارنة مع القاموس، هذا القاموس يحتوي على درجات الراي لكل كلمة. قاعدة البيانات تحتوي كذلك على اسعار العملات المتداولة والتي لها تأثير على تقلبات سعر الذهب منها (سعر الدولار الامريكي، سعر اليورو سعر الذهب الحقيقي) والتي تم تجميعها من مجموعات تداول العملات النقدية من تطبيق الفايبر في السوق الليبية الموازية. نتيجة التدريب حصول على 26 عينة مكونة من أربع متغيرات (سعر الذهب، سعر اليورو، سعر الدولار الامريكي، نتيجة الرأي). يتم استخدام نموذج دعم الاتجاه للتدريب والتنبؤ بسعر الذهب المستقبلي. تم استخدام متوسط نسبة الخطأ المطلق (MAPE) لتقييم أداء النموذج. يعتبر أداء النموذج المقترح مقارنة بالنماذج السابقة جيدا وبأقل معدل خطأ بنسبة (MAPE 1.6%).**

الكلمات المفتاحية--- **SVM، تحليل الآراء، MAPE، منصة-X، سعر الذهب.**

1. مقدمة

يعتبر التنبؤ بتقلبات أسعار الذهب أحد الاهتمامات الرئيسية في الأسواق المالية ويشكل مصدر قلق في الأسواق المالية .ولأن الذهب أظهر قدرة خاصة على تهدئة تقلبات التضخم، فإن معظم المحافظين يستخدمونه كأداة للتحكم في الأسعار. إضافة إلى ذلك، فإن التدفقات النقدية في مشاريع التعدين متقلبة وتتأثر بشكل كبير بتقلبات أسعار السلع المعدنية. ولذلك، يتم اتخاذ قرارات قبول ورفض المشاريع بناءً على توقعات الأسعار المستقبلية. وبالتالي، فإن تطوير نموذج قرار تنبؤي دقيق وقوي أمر بالغ الأهمية بالنسبة للمستثمرين. كذلك، فإن المزيد من المعلومات حول اتجاهات أسعار الذهب المستقبلية ستساعد في اتخاذ قرارات حازمة [1] . يشكل سعر الذهب أهمية في الأنظمة الاقتصادية والنقدية، فهو أداة نقدية غير شائعة في الاسواق المالية، لذلك تطور الاستثمار في الذهب مع مرور الوقت إلى الطرق التقليدية أو من خلال الاستثمار في صناديق التداول في بورصة الذهب (ETF) [2]. ولأهمية هذا المجال فقد تم تطوير العديد من النماذج وهناك العديد من الدراسات التي تناولت التنبؤ بأسعار المعادن والذهب، ولكن المجال لم يغلق بعد ويعتبر من المواضيع البحثية التي تحتاج إلى مزيد من الجهد من أجل الحصول على نموذج أكثر كفاءة ودقة. تعتمد النماذج التقليدية على سعر الذهب التاريخي (وهو دراسة اتجاهات أسعار الذهب في السنوات الماضية) [3]. بالإضافة إلى اعتماد بعض النماذج على المؤشرات الاقتصادية (وهي الدور الموجه لعوامل أسعار الذهب الكلية في سعر الذهب من مجموعة البيانات المحدودة).

واجهت النماذج التقليدية السابقة عدة تحديات، يمكن القول إن أهمها هو افتقار هذه النماذج إلى الأحداث التي لها تأثير كبير على تقلبات أسعار الذهب والمعادن مثل الكوارث الطبيعية والحروب. ويمكن العثور على هذه الأحداث على شكل أخبار على مواقع الإنترنت وصفحات التواصل الاجتماعي. تم تطوير هذه المواقع على مراحل، ومن خصائص تطوير الويب الاخير التي بدأت في الظهور والاستخدام هي تحليل الآراء، والذي يسمح للمستخدم بالبحث من خلال المؤشرات العاطفية [1]، حيث يتميز البحث من خلال تحليل الآراء بإضافة ميزات إلى طريقة البحث داخل المواقع الإلكترونية لا توجد في البحث التقليدي، ومنها أنه يمكن للمستخدم البحث من خلال الكلمات الطبيعية التي يستخدمها في حياته اليومية، وهذه الكلمات غالباً ما تكون حاضرة في آراء وتعليقات المستخدمين للموضوع الذي يبحث في هذه الفترة، يعد المحتوى الناتج عن المستخدمين واستخراج الآراء معلومات قيمة يمكن استخدامها في عمليات البحث [4][5] . يمكن للعديد من المستخدمين أيضًا تقييم نتائج البحث وفقًا للنتائج الأكثر صلة وملاءمة لإدخالات البحث [6].

من أجل التنبؤ بسعر الذهب المستقبلي، يستخدم نموذج التنبؤ عوامل لها تأثير كبير على تحديد أسعار الذهب. تم تطوير وتنفيذ عدة طرق للتنبؤ بسعر الذهب. يمكن تصنيف طرق التنبؤ إلى ثلاث طرق رئيسية [3]:

أ- النموذج الرياضي التقليدي.
ب- الذكاء الاصطناعي (AI).
ت- النماذج الهجينة.

توثق معظم الدراسات الحاجة إلى إدارة أفضل لمبيعات الذهب والاستثمار لتقليل قيمة المخاطر. هناك حاجة إلى نموذج دقيق للتنبؤ بأسعار الذهب لإظهار اتجاه تغيرات الأسعار في العقود الآجلة لإجراء التبادلات المناسبة. علاوة على ذلك، من الصعب جدًا الحصول على دالة قوية باستخدام النموذج الرياضي التقليدي [7].

1.1 الأعمال ذات الصلة

هناك العديد من الدراسات المتعلقة بالتنبؤ بأسعار الذهب على سبيل المثال، توفر الدراسة [2] العقدة الزمنية ARIMA للتنبؤ بسعر الذهب المستقبلي في متصفح البيانات التالي من نوفمبر 2003 إلى يناير 2014 للتخفيف من شراء الذهب ARIMA يحتوي على أربع خطوات رئيسية في بناء النموذج -تحديد الهوية والتشخيص والتنبؤ. من خلال هذه الخطوات يتم تحديد معلمات النموذج المؤقت الأول، ثم يتم تحديد المعلمة ومعرفة النموذج في الحال، بالإضافة إلى الخطوات التالية التحقق من

استلمت الورقة بالكامل في 02 مايو 2024 وروجعت في 10 مايو 2024 وقبلت للنشر في 15 مايو 2024
ونشرت ومتاحة على الشبكة العنكبوتية في 08 أغسطس 2024.

صحة النموذج. وأخيرا، يتم استخدام الإحصائيات الصغيرة وفترات الثقة بصحة التنبؤ وتتبع أداء النموذج. هناك بعض القيود في التنبؤ بالبيانات باستخدام نموذج ARIMA. في حالة التغيير المفاجئ، في مجموعة البيانات (عندما يكون الفرق كبير)، سيكون من الصعب التقاط التغير الدقيق، وبالتالي يصبح هذا نموذجًا غير فعال للتنبؤ في حالة البيانات غير المشتركة، ولا يوجد دليل على أن سعر الذهب خطي بطبيعته.

اقترح المؤلفون في [3] نموذج BAT-Neural (BNN) يعتمد على الشبكات العصبية الاصطناعية. تم تجهيز الشبكة الذكية المقترحة بخوارزمية BAT لجعل ANN قادرة على متابعة التقلبات. تمت مقارنة النموذج المصمم بمجموعة من النماذج مثل المتوسط المتحرك المتكامل التلقائي (ARIMA)، نظام الاستدلال العصبي الضبابي التكيفي متعدد الطبقات الإدراكية (ANFIS)، الشبكات العصبية الاصطناعية (ANN). ولتقييم أداء النموذج تم استخدام تم استخدام نموذج الجذر التربيعي للخطأ (RMSE) كمؤشر للخطأ. ويبين أن النموذج المقترح يتكون من ثلاث مراحل مختلفة. توفر المرحلة الأولى، والتي تسمى مرحلة إنشاء البيانات التعريفية، بيانات تحتوي على ميزات مفيدة لمساعدة نموذج التنبؤ. تركز المرحلة الثانية على إنتاج نموذج تنبؤ تكيفي قوي يعتمد على ANN. لتوليد الوزن تم استخدام خوارزمية الخفافيش للمساعدة في عملية التنبؤ في اختيار أفضل الأوزان وبالتالي إنتاج أفضل النتائج.

وفي الدراسة [8] اقترحت إخضاع أخبار الذهب العالمية اليومية عبر الإنترنت لنهج التنقيب القائم على النص. يتم إنشاء درجة الرأي من خلال التأكد من قطبية الرأي والكلمات في أخبار الذهب اليومية. بعد ذلك، يتم استخدام انحدار ناقل دعم المربع الأصغر (LSSVR) المحسّن بواسطة الخوارزمية الجينية (GA) لتدريب سعر الذهب والتنبؤ به في المستقبل. تم استخدام متوسط نسبة الخطأ المطلق (MAPE) لتقييم أداء النموذج. أظهرت نتائج التجربة أن مؤشر الإدخال، درجة الرأي، يمكن أن يحسن إمكانية التنبؤ بنموذج GA-LSSVR بدلالة MAPE.

2.1 تحليل الآراء

بدأ مصطلح تحليل الآراء في الظهور عام 2003. وهو مجال الدراسة الذي يحلل آراء الناس ومشاعرهم وتقييماتهم ومواقفهم وعواطفهم تجاه كيانات مثل المنتجات والخدمات والمنظمات والأفراد والقضايا والأحداث والموضوعات [5]. وهي من أحدث الأساليب التي تجري حاليًا مع تطور وانتشار مواقع التواصل الاجتماعي، ومع تزايد استخدام هذه المواقع، أصبح لدينا عدد هائل من الآراء التي يمكن الاستفادة منها في استخلاص الأنماط وتحليلها. ويتم تحليل الآراء بثلاث طرق:

أ- مستوى الوثيقة.
ب- مستوى الجملة.
ت- مستوى الجانب.

التحليل على مستوى الوثيقة يتم التصنيف حسب الشعور العام ما إذا كانت إيجابية أو سلبية أو محايدة، ويفترض ان تحتوي الوثيقة على راي واحد، واما مستوى الجملة يتم تصنيف المشاعر والتعبير عنها في الجملة بأكملها من حيث كونها إيجابية أو سلبية أو محايدة. أما التحليل على مستوى الجانب فيعتمد على استخلاص المشاعر حول سمات محددة داخل النص [6] .

يستخدم تحليل الآراء طريقتين مختلفتين. تُستخدم تقنيات التعلم الآلي لتدريب الخوارزميات على تصنيف النص. يتم تصنيف هذه الخوارزميات على أنها تعلم خاضع للإشراف أو غير خاضع للإشراف. تستخدم خوارزميات التعلم الخاضع للإشراف البيانات التي تم تصنيفها مسبقًا على أنها إيجابية أو سلبية أو محايدة. النوع الآخر من خوارزميات المشاعر يتكون من أساليب تعتمد على المعجم أو القواعد. تستخدم هذه الأساليب كلمات الرأي لنقل درجة إيجابية أو سلبية من الرأي. ويتم حساب درجة المشاعر من خلال قطبية كلمات الرأي وتضمينها [9].

والمواقف والعواطف الإيجابية أو السلبية. على سبيل المثال، يمكن

1.2.1 نظرة عامة على تحليل الآراء

تصنيف تحليل الآراء هو عملية تحديد المشاعر من نص معين إلى إحدى الفئات المحددة مسبقًا (إيجابية أو سلبية أو محايدة). يتكون تصنيف تحليل الآراء بشكل عام من أربع خطوات رئيسية كما هو موضح في الشكل (1)

شكل 1.عملية تصنيف تحليل الآراء

أ- جمع البيانات (Data Collection) : الخطوة الأولى في تحليل الآراء هي جمع البيانات المراد تحليلها. يمكن القيام بذلك عن طريق تحديد بعض الكلمات الرئيسية محل الاهتمام المراد فحصها واسترداد كل النصوص ذات الصلة بتلك الكلمات الرئيسية باستخدام بعض التطبيقات مثل Twitter API.

ب- المعالجة القبلية (Pre-processing) : تقوم بتنظيف البيانات بعد جمعها، ويحدث ذلك في عدة خطوات، وفي كل خطوة لها أهميتها في جعل البيانات تحمل فقط المعلومات المهمة التي ستكون دخلت في النموذج. وبالتالي، في هذه المرحلة، يتم تنظيف البيانات من المعلومات المكررة، وخاصة في البيانات التي تم جمعها من توتير، لذلك يجب إزالتها. وكذلك الروابط وتتم إزالة علامات التصنيف التي تم تنزيلها مع التغريدات لأنها لا تحمل أيًا منها أهمية أو مشاعر. وكذلك علامات الترقيم أو الرموز التعبيرية كما يجب تنظيف البيانات من كلمات التوقف وهي الكلمات الشائعة المستخدمة بشكل متكرر. وجودهم أو غيابهم لا يؤثر على المعنى من الجملة. وهي كلمات لا علاقة لها بالمشاعر، فتحذف، لن تؤثر على تحليل الآراء ولكنها ستسرع عملية التحليل. أي أنه يقلل من مساحة الفهرس. أمثلة على كلمات التوقف" في – من – إلى – كل – نحن ".

ت- اختيار الميزة(Feature Selection) : يهتم هذا الجزء باستخراج السمات من مجموعة المعطيات واستكشافها لمعرفة السمات الأكثر أهمية منها، ثم تعالج السمات المنتقاة لاستخدامها في النموذج.

ث- تصنيف المشاعر (Sentiment Classification) : هي الخطوة الأخيرة والحاسمة، حيث يتم تصنيف النص، على سبيل المثال، سلبي ايجابي محايد. يمكن استخدام ثلاثة أنواع في تصنيف المشاعر: النهج القائم على المعجم، والنهج القائم على التعلم، والمنهج الهجين.

2.2.1 التنقيب على الرأي

ويمكن تعريفها بأنها العمليات التي تتم على البيانات غير المنظمة ومن ثم تصنيفها .هناك ثلاث مهام لعملية التنقيب على الرأي، وهي تحديد ما إذا كان النص يعبر عن شيء إيجابي أو سلبي أو محايد، ثانيا، عملية استخلاص الرأي وأخيرا اكتشاف الرأي وتلخيصه [10].

3.2.1 معجم المشاعر (Sentiment Lexicons)

ويسمى أيضًا معجم الرأي وقاموس المشاعر. يشير إلى قاموس الكلمات أو العبارات أو التعابير. يمكن استخدام هذه المصطلحات لحساب درجة المشاعر لكل مستند ثم تصنيف المستندات حسب قطبيتها العالية. تعتبر هذه الكلمات بمثابة مؤشرات جيدة لتحليل الآراء. يطلق عليها في الأدبيات كلمات المشاعر، أو كلمات الرأي، أو الكلمات القطبية، أو الكلمات التي تحمل الرأي. هناك كلمات شائعة يمكن استخدامها للتعبير عن المشاعر

استخدام الجيد والممتاز والرائع للتعبير عن الشعور الإيجابي بينما يتم استخدام السيئ والفظيع للإشارة إلى الشعور السلبي [11]. يمكن استخدام المعجم بطرق مختلفة في تصنيف تحليل الآراء. على سبيل المثال على سبيل المثال، يتم حساب درجة المشاعر لكل مستند من خلال جمع

درجات المعجم لكل مصطلح في المستند مع الأخذ في الاعتبار حالات النفي والحالات الخاصة [12]. إذا كانت النتيجة النهائية أكبر من الصفر، فيمكن تصنيف الوثيقة على أنها إيجابية، وإلا فهي سلبية [13].

4.2.1 التحديات في تحليل الآراء العربية

تكمن الصعوبة الرئيسية عند إجراء تحليل الآراء في وسائل التواصل الاجتماعي العربية في حقيقة أن التواصل في سياق وسائل التواصل الاجتماعي يتم باستخدام اللغة العربية العامية أو الجدلية بدلاً من اللغة العربية الفصحى الحديثة الأكثر رسمية (MSA). تختلف مفردات اللغة العربية الجدلية عن اللغة العربية الفصحى. علاوة على ذلك، فإن بنية الجملة أكثر عشوائية، وهذا هو السبب في أن مهمة تحليل هذا النص صعبة للغاية [14].

3.1 نموذج دعم المتجهات SVM

SVM عبارة عن خوارزمية تعلم آلي بسيطة جدًا لكنها قوية يمكن استخدامها للتصنيف وكذلك الانحدار على الرغم من استخدامها بشكل شائع للتصنيف [15]. إنها تؤدي أداءً جيدًا في مجموعات البيانات الصغيرة والمتوسطة الحجم ويسهل ضبطها للغاية [16]. يتم استخدام مصنف دعم المتجهات كنموذج أساسي. أظهرت الأبحاث السابقة أنه يمكن استخدام نموذج دعم المتجهات للتنبؤ بتحركات الأسعار في الأسواق المالية [17]. من السهل تنفيذ نموذج دعم المتجهات باستخدام الحزم الموجودة في لغة R. تم استخدام المصنف SVM في هذه الدراسة. هناك ثلاث مراحل أساسية لتحليل SVM:

أ- اختيار الميزة
ب- تدريب واختبار المصنف
ت- تقييم الأداء تجدر الإشارة إلى أن هذه المراحل ليست خاصة بـ SVM ولكنها موجودة في معظم أساليب التعلم الآلي.

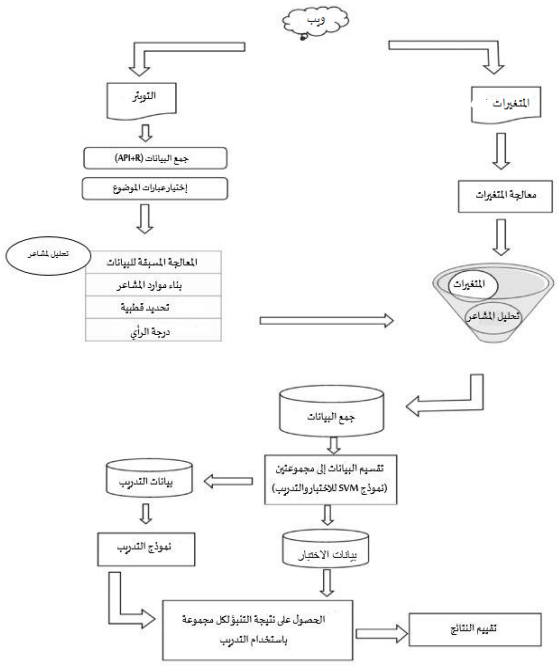
2. منهجية البحث

منهجية البحث لدينا تتكون من خمس مراحل في المرحلة الأولى؛ تم جمع البيانات من منصة -X. في المرحلة الثانية، يتم تطبيق العديد من مهام المعالجة المسبقة. وفي المرحلة الثالثة تم استخدام قاموس AraSenti-الذي طوره باحثون لإضفاء العواطف على التغريدات. في المرحلة الرابعة، يتم إنشاء الميزات التي سيتم إدخالها في وحدة SVM. المرحلة الاخيرة تم تنفيذ نموذج SVM.

1.2 جمع البيانات لتحليل الآراء

في هذا القسم، نصف مرحلة جمع البيانات. استخدمنا منصة-X كمصدر لمجموعة البيانات الخاصة بنا. منصة -X هو موقع للتواصل الاجتماعي حيث يمكن للأشخاص مشاركة مشاعرهم وآرائهم وأفكارهم مع الآخرين. لقد اخترنا منصة -X كمصدر بياناتنا لعدة أسباب:

يوفر منصة-X ويدعم ثلاثة مستويات للوصول إلى البيانات من خلال ثلاث حزم تسمح للباحثين أو المستخدمين بجمع البيانات. هذه الحزم هي الحزم القياسية، والمميزة، والمؤسسية. كما يوفر أيضًا ميزات لجمع التغريدات في الوقت الفعلي. الحزمة القياسية مجانية وتسمح للمستخدم بالبحث عن التغريدات من آخر 7 أيام. يمكن استخدام الحزمة المميزة



شكل 2. المخطط الانسيابي لنموذج التنبؤ بسعر الذهب بناءً على آلة ناقل الدعم (SVM)وتحليل الآراء.

بعقد رسمي وموافقة حيث يسمح منصة-X للمستخدم بالبحث عن التغريدات في آخر 30 يومًا أو الأرشيف الكامل منذ عام 2004.

لتخزين التغريدات المجمعة، استخدمنا جدول بيانات Excel. يشير Excel إلى عدم وجود قاعدة بيانات علائقية. ولذلك، يتم تجنب مقدار الحمل لقاعدة بيانات علائقية. وهذا يوفر تخزين سريع للبيانات. لدى Excel تأثير إيجابي على البيانات الضخمة والتطبيقات في الوقت الحقيقي.

يوفر Excel محركًا للاستعلام بسهولة عن التغريدات المجمعة باستخدام معايير مختلفة. فهو يوفر واجهات برمجة تطبيقات Java وPhyton وR للوصول إلى البيانات المجمعة حيث يمكننا استخدامها بسهولة عندما نحتاج إلى استرداد البيانات للمعالجة المسبقة والتعلم الآلي.

إن واجهات برمجة التطبيقات المتدفقة لنموذج الكائن-X هي كائن JSON. يحتوي على قيمة المفتاح، حيث يمثل المفتاح اسم الحقل، وتمثل القيمة بيانات الحقل. يوضح الشكل (3) مثالاً لإحدى التغريدات المستلمة باستخدام tweets API.

وتحتوي على عدة سمات مثل نص التغريدة، ووقت الإنشاء، ومعلومات المستخدم، ولغة التغريدة، وإعادة لمن، والاقتطاع أم لا، وإعادة التغريد أم لا، والعديد من المقاييس الأخرى. يساعد عند استخراج نص التغريدات.

created_at	screen_name	text
44621.73	218Tv	# الحرب # روسيه # اوكرانيه والازمه ليبيه # ليبيا
44621.67	218Tv	# تجار # الدقيق يستغلون حرب # اوكرانيا لاشعال الرغيف اليبي
44621.63	218Tv	سي يستقيل من منصبه احتجاجا على الهجوم الروسي على # اوكرانيا

شكل 3. مثال لأحد الكائنات المستلمة من واجهات برمجة تطبيقات X

2.2 المعالجة المسبقة

قنوات إعلامية عربية (محلية أو أجنبية)، تقدم معلومات وأخبار عالمية ومحلية (ليبية) تقدم الأخبار باللغة العربية الفصحى، تحمل هذه

التغريدات مشاعر متوقعة في تذبذب سعر الذهب في السوق المحلية. تم طلب التغريدات من القنوات الإعلامية المذكورة في الجدول (1) والجدول (2).

جدول 1. اسماء القنوات المحلية

	اسم القناة
1	TV218
2	Alwasatnewsly
3	EanLibya
4	Fawaselmedia
5	Libyaalahrartv
6	LibyasChannel

جدول 2 القنوات الغير محلية

	أسماء القنوات
1	AJArabic
2	AlArabiya_Brk
3	BBCArabic
4	Skynewsarabia

توفر منصة X- أدوات تساعد الباحث على التعامل مع البيانات من جلب ومعالجة مسبقة وتطبيع وكل ما يحتاجه الباحث لتنظيف البيانات قبل عملية التنبؤ وذلك باستخدام API كما هو موضح في الشكل (3) لذلك اخترنا منصة X- كمصدر بيانات التنقيب عن النصوص.

1.2.2 تنظيف البيانات (Data cleaning)

هذه هي المرحلة الأولى لتحليل الآراء بعد عملية جمع البيانات من منصة-X. تساعد هذه المرحلة على تقليل البيانات المزعجة وترتيب البيانات بطريقة تسهل تحليلها. وللقيام بذلك، استخدم وظائف جاهزة لتنظيف البيانات.

استخدام بعض الدوال الجاهزة لإزالة العديد من البيانات غير الضرورية وغير المرغوب فيها التي لا تحتوي على أي معنى للمشاعر، على سبيل المثال: محددات مواقع (URL)، وبعض الرموز مثل({ } \ $ & # $ %). بالإضافة إلى ذلك تتم إزالة الحروف المتكررة بحرف واحد، و إزالة علامات الترقيم أيضًا(، : ”)، وكذلك حذف الأرقام العربية و الأجنبة و اليونانية. وإزالة الأحرف الاجنبية او الغير عربية، والتشكيل.

2.2.2 احتفظ بالكلمات الرئيسية (Keep keywords)

هي الاحتفاظ بالكلمات الرئيسية أثناء عملية تنظيف البيانات، تلك الكلمات التي يتم تعديلها بشكل غير صحيح عن طريق حذف أو إضافة أحرف تغير معنى الكلمة. لذلك يتم الاحتفاظ بها دون التغير فيها.

3.2.2 التغريدات غير المرغوب فيها (Spam Tweets)

عند جمع التغريدات نحتاج إلى فرزها ضمن التغريدات التي تهم موضوعنا، ويتم تمييز التغريدات حسب الكلمات المفتاحية، بحيث يتم حذف أي تغريدة تحتوي على كلمة (كلمات غير مفتاحية او غير مرغوب فيها) يتم تخزين الكلمات في مصفوفة (Spam Tweets).

4.2.2 كلمات التوقف (Stop Words)

تعتبر كلمات التوقف خالية من المشاعر والتي لها تأثير أقل على المعني مثل (من، في، على)، وبالتالي يتم حذفها من التغريدات.

السلبية. يمكن أن يكون لكلمات الرأي هذه قطبية إيجابية (على سبيل المثال، " شكر" أو " مدح ")، أو يمكن أن تعبر عن دلالة سلبية (لـ على سبيل المثال، "منزعج" أو "أسفل"). بالإضافة إلى ذلك، إدارة حدوث النفي أمر بالغ الأهمية. ورغم أن مصطلح "متفائل"، على سبيل المثال، إيجابي بشكل واضح، فإن مصطلح "غير متفائل" سلبي رغم احتوائه

3.2 بناء موارد المشاعر (Building Sentiment Resource)

يتم في هذه المرحلة إعداد الآراء لصورتها النهائية وتحويل مجموعات البيانات إلى نماذجها النهائية. استخدمنا في هذه الخطوة طريقة المعجم واستخدمنا معجم الرأي تحديدًا لتصنيف درجة جميع التغريدات.

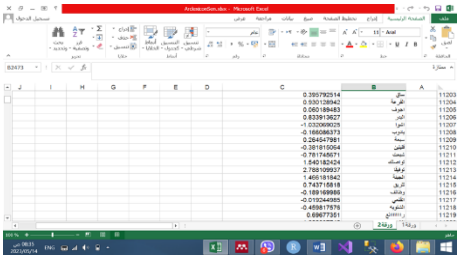
1.3.2 بناء معجم آراسنتي (Building AraSenti Lexicon)

هذا المعجم عبارة عن قائمة بالمصطلحات العربية وارتباطاتها بالمشاعر. ويظهر الارتباط في الشكل (4). إذا كانت نتيجة x1 أكبر من نتيجة x2، فإن x1 يعتبر أكثر إيجابية من x2. تم استخراج الكلمات الموجودة في المعجم من مجموعة بيانات كبيرة من التغريدات العربية.

تم استخدام طريقتين لتوليد المعجم:

أ- منهج الترجمة، حيث تمت مقارنة الكلمات العربية المترجمة إلى الإنجليزية مع المعاجم العاطفية الإنجليزية. تم تعيين درجة المشاعر للكلمات الإيجابية وتم تعيين درجة المشاعر للكلمات السلبية -1.

ب- نهج المعلومات المتبادلة حيث يتم حساب درجة المشاعر باستخدام المعلومات المتبادلة (PMI) بين المصطلح والفئات الإيجابية والسلبية كما مذكور في المعادلة (1):



شكل 4. معجم المصطلحات العربية مع درجة المشاعر

$$SenScor = PMI(w, pos)? \ PMI(w, neg) \quad (1)$$

حيث w هو مصطلح في المعجم. PMI (w, pos) هي درجة PMI بين w والطبقة الإيجابية، وPMI (w, neg) هي درجة PMI بين w والطبقة السلبية. تشير الدرجات الأكبر من 0 إلى أن هذا المصطلح لديه ميل أكبر للتزامن مع البذور الإيجابية مقارنة بالبذور السلبية --- وبالتالي من المحتمل أن يكون إيجابيًا. تشير الدرجات الأقل من 0 إلى أن هذا المصطلح لديه ميل أكبر للتزامن مع البذور السلبية مقارنة بالبذور الإيجابية وبالتالي من المحتمل أن تكون سلبية. ثم تم دمج المعجمين وإزالة النسخ المكررة. يحتوي المعجم على 225332 كلمة، ويعتبر من القواميس الجيدة في مجال استطلاع الرأي، حيث يحتوي على عدد كبير من الكلمات السلبية والإيجابية، والتي من دورها كلما زاد التحسن في مستوى النتائج.

4.2 تحديد قطبية الرأي (Opinion polarity identification)

بمجرد تحديد البيانات التي تم التعليق عليها، يكون الهدف هو تحديد الآراء المعبر عنها في هذه البيانات؛ على وجه التحديد، إذا تم ذكر السمة سلبا أو إيجابا. وللقيام بذلك، من الضروري تحديد ما يسمى بالتعبيرات الضمنية عن الرأي. يتم جمع كافة الجمل التي تحتوي إما على عبارة أو معلومات مبنية على الرأي والتي تعبر عن آراء المستخدمين الإيجابية أو

على كلمة إشارة إيجابية. ولتجنب مثل هذه التفسيرات الخاطئة، تم التعامل مع كل كلمة في المعجم على حدة، حيث أن كلمة "متفائل" لها معنى إيجابي ، في حين أن كلمة "غير متفائلة" هي كلمة سلبية لها وزن آخر. علاوة على ذلك، يجب أن نبحث عن درجة الكلمة التي تسبق كلمة الرأي أو تليها؛ تحتوي كل كلمة نتيجة متطابقة على قيمة قوة محددة

مسبقًا تُستخدم لحساب قوة المشاعر. يحتوي قاموس AraSenti على 225332 كلمة بدرجة. بالإضافة إلى ذلك، كل كلمة ذات نتيجة متطابقة لها قيمة قوة محددة مسبقًا يتم استخدامها لحساب الشعور بقوة الكلمة.

يتم احتساب درجة الرأي من الأخبار المؤثرة على الذهب وفق الخطوات التالية:

الخطوة 1. تتم عملية مقارنة العبارات والمعجم. إذا كانت الكلمة موجودة، انتقل إلى الخطوة 2.

الخطوة 2. إذا تم العثور على كلمة "عبارة موضوع" + "استقطاب الرأي"، انتقل إلى الخطوة 3، أو انتقل للتحقق من الرسالة التالية.

الخطوة 3. أخذ كلمة الرأي المحددة، انتقل ذهابًا وإيابًا للتحقق من الكلمة لتحديد ما إذا كانت درجة الكلمة موجودة في قاموس AraSenti. إذا وجدت، يتم حساب قوة المشاعر بناءً على النتيجة (أي الوزن) المحددة مسبقًا في AraSenti.

5.2 درجة الرأي (Opinion score)

للحصول على قيمة بيانات الموضوع يتم تطبيق المعادلة (2). يتم احتساب درجة الرأي على أساس قطبية كلمات الرأي المحددة المرتبطة بعبارات الموضوع ووزن كلمات النتيجة إن وجدت. يتم تعريف وظيفة التسجيل على النحو التالي:

$$\textbf{OpionScore j} = \sum_{i=1}^{n}(w) \qquad (2)$$

OpionScore j: درجة الرأي لعبارات الموضوع j

w : وزن الكلمة من الدرجة الأولى في AraSenti

n : عدد كلمات الرأي المحددة في التغريدات

إذا كانت عبارة الموضوع "أ" حاصلة على درجة رأي إيجابية، فسيتم تفسيرها على أنها مناقشة إيجابية للسمة. وكذلك إذا كان هذا المجموع سالبا، فإنه يفسر على أنه مناقشة السمة بطريقة سلبية. نقاط الرأي يتم احتساب نقاط الرأي اليومية للأخبار اليومية من يناير 2021 إلى ديسمبر 2022 بناءً على المعادلة (2) ويضاف المبلغ حسب المعاملة، كما هو موضح في الجدول 1، ويتم تسويته مع متوسط سعر الذهب الفعلي شهريًا باستخدام المعادلة (3).

$$Normalized\ value = \frac{\textbf{Original value - Min value}}{Max\ value - Min\ value} \qquad (3)$$

6.2 مجموعة البيانات من البيانات المتغيرة (Data set of (Variable Data

تتكون مجموعة البيانات لدينا من مصدرين، الأول من موقع-X كما ذكرنا في الجزء السابق، والجزء الثاني يتم جمعه من قبل مجموعة فايبر التي تنشر الأسعار بشكل دوري. تم جمع الأسعار بشكل دوري من تاريخ يناير 2021 إلى تاريخ ديسمبر 2022 من برنامج فايبر الخاص بالغرفة الذي يقوم بنشر الأسعار للسوق الموازي بشكل منتظم مما يعرف بـ (سوق المشير) في مدينة طرابلس. تم جمع البيانات عبر مجموعة من الخطوات:

أ- جمع الأسعار في ملف Excel.

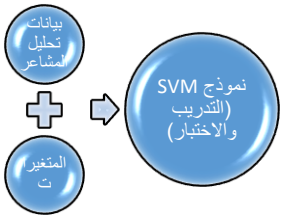
ب- باستخدام الدالة الجاهزة في Excel، تم فصل كل أسبوعين لتشكيل واحدة منفصلة.

ت- احسب الوسط الحسابي لكل اسبوعين.

ث- بعد الانتهاء من جمع البيانات ومعالجتها، يتم جمع كافة المتغيرات في جدول واحد. يحتوي هذا الجدول على بيانات الإدخال لنموذج SVM.

7.2 جمع البيانات لنموذج التدريب

بعد أن يتم جمع بيانات أسعار العملات والذهب، وتحليل الرأي لبيانات موقع-X، يتم تجميع المخرجات في ملف خاص يحتوي على البيانات (متغيرات النموذج). ويوضح الشكل (5) جمع البيانات لتدريب واختبار نموذج SVM. تنقسم هذه البيانات إلى قسمين: بيانات التدريب وبيانات الاختبار.



شكل 5. جمع البيانات لتدريب واختبار نموذج SVM

8.2 النموذج المقترح بناءً على SVM وتحليل الآراء

تم تصميم نموذج التنبؤ بسعر الذهب بناءً على تحليلات المشاعر وSVM، أولاً وقبل كل شيء، يتم جمع الأخبار المحلية والعالمية بناءً على الكلمات الرئيسية التي لها علاقة مباشرة بالأحداث الأكثر تأثيرًا على سعر الذهب (مثل الحروب والمشاكل الاقتصادية وما إلى ذلك) ومن ثم معالجتها مسبقًا لإجراء تحليل الآراء والحصول على درجة الرأي لكل مجموعة من التغريدات (كل أسبوعين). وفي هذه الأثناء، يتم جمع وتحليل المتغيرات المستخدمة بشكل متكرر، مثل المتغيرات الاقتصادية والمعادن الثمينة للعثور على المتنبئين ذوي الصلة. يتم دمج نتائج الرأي والتنبؤ المختارة وتقسيمها إلى بيانات تدريب واختبار، والتي يتم استخدامها بعد ذلك في النموذج المقترح للحصول على نتيجة التنبؤ لكل مجموعة معاملات. بعد جمع البيانات من تحليل التوجهات وأسعار العملات والسعر الحقيقي للذهب من السوق الموازية في ملف اكسل، تنقسم البيانات إلى قسمين: الجزء الأول للتدريب بنسبة 80%، والجزء الثاني للتدريب الاختبار بنسبة 20%. كما هو موضح في الجزء التالي (التنفيذ).

9.2 نموذج تقيم الأداء

تم اقتراح العديد من نماذج القياس واستخدامها لتقييم دقة التنبؤ للنماذج مثل MAPE وMSE وما إلى ذلك. والمقياس الأكثر استخدامًا هو MAPE. من المزايا المهمة لهذا الإجراء أنه لا يعتمد على الأحجام المتغيرة المتوقعة. MAPE هو مقياس الخطأ الأكثر ملاءمة لتقييم أداء التنبؤ. في هذه الدراسة، تم استخدام MAPE لتقييم دقة التنبؤ من خلال التحليل المقطعي.

$$\text{MAPE} = 100\frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right| \qquad \textbf{(4)}$$

كما توضح المعادلة (4): حيث At هي القيم الفعلية للفترة t، و Ft القيمة المتوقعة للفترة t، و n عدد عينات التدريب. كلما كان MAPE أصغر، كلما كانت القيم المتوقعة أقرب إلى البيانات التاريخية الفعلية، وكان نموذج التنبؤ أكثر دقة.

تم اختبار أداء النموذج المقترح بواسطة (MAPE) ومقارنة نتيجة النموذج المقترح مع أداء النماذج الأخرى التي تم ذكرها في الفقرة السابقة (مراجعة الأدبيات).

3. النتائج

نموذج المقترح الذي تم تدريبه حقق أداءً جيدًا في الاختبار والتدريب. يعتبر أداء النموذج مقارنة بالنماذج السابقة جيد جداً عند قياس الأداء باستخدام نموذج الخطأ المطلق. بشكل عام، النموذج يعمل بشكل جيد. بالنظر إلى متوسط نسبة الخطأ المطلق (MAPE)، يتوقع النموذج تقريبًا، أن أي حركة في السعر الحقيقي للذهب هي حركة إيجابية بشكل خاص في درجة الرأي.

1.3 نتائج سمات النموذج Results of features modes

يعرض هذا الجزء نتائج متغيرات النموذج والتي تم الحصول عليها من تحليل البيانات من أكثر من 102909 تغريدة من القنوات المحلية والأجنبية كما ذكرنا في الفقرة السابقة، ومتوسط أسعار العملات والتي سيتم شرحها على شكل جداول والرسوم البيانية. تم جمع أسعار الذهب من تطبيق فايبر (مجموعات التداول العملة) وتم حساب المتوسط لكل أسبوعين. ويوضح الشكل (6) متوسط أسعار الذهب الفعلية مع التاريخ لكل اسبوعين، ويوضح تقلبات أسعار الذهب من شهر نوفمبر 2021 إلى نوفمبر 2022.

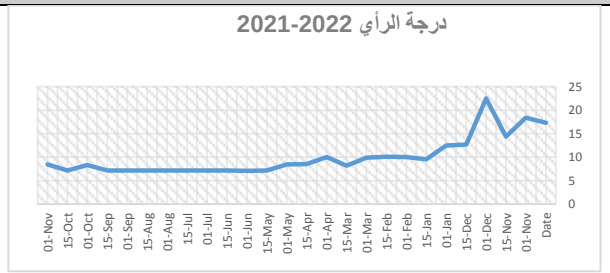
شكل 6. يوضح متوسط أسعار الذهب الفعلية مع التاريخ

وبعد جمع البيانات من موقع-X والمعالجة المسبقة للبيانات وتحليل الآراء، تم الحصول على درجة الرأي كما هو موضح في الشكل (6) متوسط درجة المشاعر للتاريخ لكل أسبوعين. يُظهر التغير في نتيجة تحليل الآراء من نوفمبر 2021 إلى نوفمبر 2022.

يبين الجدول (1) أيضاً سمات مدخلات النموذج، ومتوسط الأسعار خلال فترة أسبوعين، ودرجة الرأي في الأخبار المحلية والعالمية.

معاملات الشهر	درجة الرأي	متوسط سعر الذهب الفعلي	متوسط اليورو الفعلي	متوسط الدولار الفعلي
01-Nov	17،29993365	205.142857	5.54357143	**4.9388141**
15-Nov	18،37837714	210.46875	5.56	**5.02636719**
01-Dec	14،35235075	205.666667	5.576	**4.99865385**
15-Dec	22،48076556	207.928571	5.576	**5.02539063**
01-Jan	12،67227428	209.035714	5.53016667	**4.98852778**
15-Jan	12،43830503	208.266667	5.47903846	**4.91304688**
01-Feb	9،532959007	209.1	5.48035714	**4.9165625**
15-Feb	9،983615352	218.1	5.46519231	**4.91891667**
01-Mar	10،06319672	227.2	5.3966129	**4.97483333**
15-Mar	9،868213427	222.125	5.391875	**4.94625**
01-Apr	8،201271847	224.846154	5.37159091	**4.96283333**
15-Apr	10	227.066667	5.35071429	**5.01583333**
01-May	8،493783689	225.357143	5.33785714	**5.11127976**
15-May	8،463785704	221.769231	5.33365385	**5.08089286**
01-Jun	7،154782022	220.466667	5.34395833	**5.046**
15-Jun	7،118946911	221.8	5.36666667	**5.10983333**
01-Jul	7،142857143	211.5	5.243	**5.11428571**
15-Jul	7،142857143	209	5.21428571	**5.11**
01-Aug	7،142857143	215.733333	5.18666667	**5.09666667**
15-Aug	7،142857143	213.533333	5.1575	**5.14666667**
01-Sep	7،142857143	211.533333	5.16875	**5.1725**
15-Sep	7،142857143	207.4	5.13166667	**5.225**
01-Oct	7،142857143	210.2	5.16	**5.274**
15-Oct	8،333333333	206.933333	5.18666667	**5.2575**
01-Nov	7،142857143	210.666667	5.23666667	**5.2325**
15-Nov	8،463785704	214.461538	5.33	**5.14**

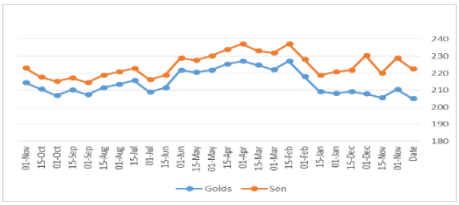
جدول 1. درجات الرأي للأخبار مع متوسط سعر الذهب



شكل 7. يوضح متوسط درجة الشعور بالتاريخ 2021-2022

ويبين الشكل (8) رسماً بيانياً لكل من متوسط السعر الحقيقي للذهب مع قطبية الآراء للأخبار. أظهرت النتائج أن متوسط الذهب الفعلي يتأثر بمعدل الآراء.

ويبين الجدول (1) نتائج الرأي للأخبار مع متوسط سعر الذهب لكل أسبوعين من نوفمبر 2021 إلى نوفمبر 2022، حيث تم تقسيم الفترات كل أسبوعين. ويتبين من خلال القيم التي تم الحصول عليها أن سعر الذهب يتأثر بنتيجة تحليل الرأي، كما هو مبين في الشكل (7).



شكل 8. مقارنة درجات الرأي الشهرية وأسعار الذهب للأعوام 2021-2022.

2.3 تحليل الأداء

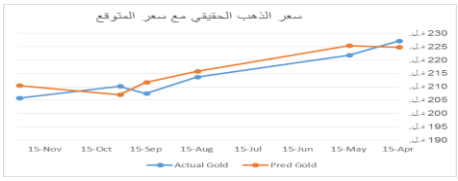
ولتحليل أداء النماذج المدروسة، تقرر حساب خطأ التنبؤ الممثل للنموذج. يتم تعريف الخطأ على أنه الفرق بين القيم الفعلية والقيم المتوقعة. النموذج ذو نسبة الخطأ الأقل يحصل على أعلى درجة من الدقة.

في هذا البحث، تم استخدام إحصائية تقارب متوسط النسبة المئوية للخطأ المطلق (MAPE) لحساب مؤشر الأداء. يتم حساب MAPE باستخدام المعادلة (2)

كما في الجدول (2) تم قياس متوسط النسبة المطلقة للخطأ (MAPE) على أسعار الذهب الفعلية مع الأسعار المتوقعة لمجموعة الاختبار، وأظهرت النتائج أن نسبة الخطأ بلغت 1.6%، وهي نتيجة مبشرة مقارنة بالنماذج السابقة.

جدول 2. سعر الذهب الفعلي وتوقعات SVM

Transaction Number	Actual Gold Price	Forecasted Gold Price
3	205.6666667	210.46875
12	227.0666667	224.8461538
14	221.7692308	225.3571429
20	213.5333333	215.7333
22	207.4	211.53
23	210.2	206.933
	MAPE 1.6%	



شكل 9. سعر الذهب الفعلي وتوقعات SVM من عام 2021 إلى عام 2022.

ويبين الشكل (9) أداء التنبؤ لنموذج SVM المقترح (الخط الأزرق) مقابل القيم الحقيقية. يقدم الرسم البياني أداءً تنبؤيًا جيداً حيث تمكن من التعامل مع اتجاهات وتقلبات أسعار الذهب.

3.3 التقييم والمقارنة

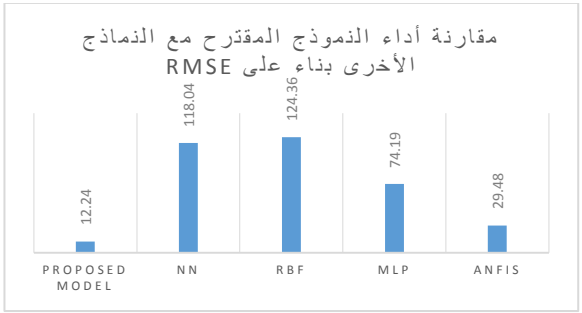
وبالمقارنة مع النماذج التي تم مناقشتها في الفصل السابق اثبت النموذج المقترح يتفوق على بقية النماذج مع أقل معدل خطأ قدره (12.24 RMSE) كما هو مبين في الجدول (3).

ملاحظة: تم استخدام (Root Mean Square Error) من اجل المقارنة مع بقية النماذج.

جدول 3 التنبؤ بالأداء للنماذج المدروسة بناء على نموذج (RMSE)

النماذج	ARIMA	ANFIS	MLP	RBF	NN	النموذج المقترح
RMSE	166.87	29.48	74.19	124.36	118.04	12.24

ملاحظة الفرق في مجموعة البيانات وفترة الدراسة (فترة جمع البيانات)



شكل 10. التنبؤ بالأداء للنماذج المدروسة بناء على (RMSE)

وكما هو مبين في الشكل (10) فإن النموذج المقترح لديه أقل معدل خطأ مقارنة بالنماذج السابقة.

4.3 المناقشة

يعتبر أداء النموذج المقترح مقارنة بالنماذج السابقة جيد جدا عند قياس الأداء باستخدام متوسط نسبة الخطأ المطلق (MAPE). تم تدريب النموذج على مجموعة بيانات تم جمعها من مصدرين (موقع- X وتطبيق Viber).

تم التواصل مع إدارة موقع-X وطلب تغريدات مجانية الخاصة بالطلبة من خلال حساب المطور الذي تم إنشاؤه. وكانت نتيجة التواصل مع إدارة موقع-X الحصول على عدد التغريدات المطلوب وهو 102000 بصيغة Excel تقريبًا. لمدة سنة تقريبا.

وتم تطبيق المعالجة المسبقة على التغريدات، مما أدى إلى خفض عدد التغريدات إلى (23601). كما تم تطبيق مفهوم التطبيع وحذف الكلمات بدون العواطف كما ذكرنا في الفصل السابق. وتقارن الكلمات مع المعجم ويتم الحصول على قيمة الشعور المذكورة في المعجم. وتم حساب المتوسط الحسابي لهذه القيم لكل أسبوعين كما هو مبين في الجدول (1). وكانت النتائج واعدة بالمقارنة مع تقلبات سعر الذهب الحقيقي. يوضح الشكل (9) مدى تناسب تقلبات سعر الذهب الحقيقي مع قيمة تحليل الآراء ، والذي بدوره يمكن القول أن تحليل الآراء يمكن أن يزيد من تحسين أداء النموذج في التنبؤ.

وتم تدريب النموذج على المتغيرات الأربعة (سعر الذهب، سعر الدولار الأمريكي، سعر اليورو، وقيمة تحليل الآراء) بنسبة 80:20 بالمائة. تم حساب متوسط نسبة الخطأ المطلق (MAPE).

الحصول على نسبة 1.6% وهي نسبة جيدة جداً مقارنة بالنماذج السابقة كما هو موضح في الجدول (2).

وبعد تطبيق النموذج المقترح والحصول على النتائج المذكورة أعلاه يمكن القول بأن النموذج المقترح بشكل عام يعمل بشكل جيد ويحقق أداء جيد. ونظراً لنموذج متوسط النسبة الخطأ المطلقة (MAPE)، فإن أداء النموذج المقترح يعتبر جيداً عند قياس ومقارنة الأداء مع النماذج الأخرى.

4. الاستنتاج

تم تطبيق النموذج المقترح الجديد على مجموعة بيانات أسعار الذهب، وأظهرت النتائج تفوق النموذج المقترح في أداء النماذج القياسية الأخرى مثل ARIMA وANN وANFIS. وبالتالي، يمكن القول أن النموذج المقترح أعطى نتائج واعدة يمكن استخدامها للتنبؤ بأسعار الذهب وكذلك الأسواق المالية الأخرى.

في الوقت الحاضر، لا يمكن لأساليب التقدير الكلاسيكية التنبؤ بدقة بتقلب أسعار السلع الأساسية والتي من بينها سعر الذهب والمعدن بسبب الأحداث المتكررة والمتغيرة بسرعة في العقد الماضي. تعتمد عملية التنبؤ في معظم الدراسات على التعرض للاتجاهات المستقبلية وتوقعاتها بناءً على البيانات التاريخية. في هذه الدراسة، تم دمج درجة الرأي التي تم الحصول عليها من استكشاف النصوص مع معدلات العمل الفعالة المختارة من المتغيرات التقليدية كتنبؤ لأسعار الذهب باستخدام نموذج SVM، دعما لاستخدام المؤشرات التاريخية التقليدية. ولذلك أظهرت نتائج الدراسة تحسنا في أداء النموذج المقترح. بالإضافة إلى ذلك، وبما أن سعر الذهب يعتمد بشكل كبير على الثقة في السوق الحالية، فإن نتائج الرأي هي اتجاهات السوق الحالية، لذلك أصبح مؤشرات فعالة. تشير نتائجنا إلى أن استخدام درجة الرأي كمؤشر مدخل لنماذج التنبؤ يمكن أن يؤدي إلى نتائج تنبؤ أكثر دقة من المتغيرات التقليدية التي تستخدم للتنبؤ بأسعار الذهب التاريخية. علاوة على ذلك، في هذه الدراسة، تم تطبيق SVM للتنبؤ بأسعار الذهب باستخدام نتائج الرأي والعملات المختارة كمتغيرات تنبؤ للنموذج المقترح. تم استخدام نموذج MAPE لتقييم اداء النموذج المقترح.

5. النطاق المستقبلي للدراسة

اقترحت هذه الدراسة نموذجًا للتنبؤ بسعر الذهب باستخدام نموذج SVM اعتمادا على تحليل الآراء للأخبار المحلية والعالمية باللغة العربية الفصحى. من الضروري إجراء مزيد من الابحاث لاختبار هذا النموذج المقترح وتطوير ما يلي:

أ- تطوير المعجم المستخدم في تحليل الرأي بحيث يركز على مصطلحات الأخبار باللغة العربية الفصحى.

ب- استخدام خوارزميات قطع الكلمات وتجذيرها (إرجاع الكلمة إلى جذرها) مع مراعاة مشكلة النفي.

ت- زيادة مدة الدراسة من سنة إلى سنتين على الأقل، وزيادة عدد التغريدات وبيانات التدريب والاختبار بهدف تحسين أداء النموذج.

المراجع

[1] Shafiee, Shahriar , Topal and Erkan, "An overview of global gold market and gold price forecasting," in *Resources Policy*, Australia-Kalgoorlie, 2010.

[2] Guha, Banhi , Bandyopadhyay and Gautam, "Gold Price Forecasting Using ARIMA Model," in *Journal of Advanced Management Science*, Durgapur,India, 2016.

[3] Hafez and Akhavan, "Forecasting Gold Price Changes: Application of an Equipped Artificial Neural Network," in *Technology Foresight Group*, Tehran, 2018.

[4] M. C. &. V. V. Smith, Smith, Matthew and Szongott, Christian and Henne, Benjamin and Von Voigt and Gabriele, "Big data privacy issues in public social media," in *IEEE International Conference on Digital Ecosystems and Technologies (DEST)*, 2012.

[5] Hu ،Minqing ، Liu and Bing" ،Mining and Summarizing Customer Reviews "،Chicago ، .2004

[6] X. Q. X. &. Z. G. Lei, Lei, Xiaojiang , Qian, Xueming , Zhao and Guoshuai, "Rating Prediction Based On Social Sentiment From Textual Reviews.," in *(2016). Rating Prediction Based On Social Sentiment From Textual Reviews.*, 2016.

[7] Ac hireko و Ansong" ،Stochastic model of mineral prices incorporating neural network and regression analysis .2000 "،

[8] Yuan ،Fong-Ching ،Lee ،Chao-Hui ،Chiu Chaochang" ،Using Market Sentiment Analysis and Genetic Algorithm-Based Least Squares Support Vector Regression to Predict Gold Prices "،*nnovation Center for BigData and Digital Convergence* ،Taiwan .2020 ،

[9] W. Medhat ،A. Hassan and H. Korashy ، "Sentiment analysis algorithms and applications: Asurvey "،*Ain Shams engineering journal* ،Cairo .2014 ،

[10] Mewari ،Ritu ، Singh ،Ajit ، Srivastava و Akash" ،Opinion Mining Techniques on Social Media Data "،*International Journal of Computer Applications* .2015 ،

[11] Liu and Bing, "Sentiment Analysis and Opinion Mining," in *Synthesis Lectures on Human Language Technologies*, Toronto, 2012.

[12] A. Amal ، K. Manal and T. Mounira" ،Arabic Sentiment Analysis Using Deep Learning and Ensemble Methods "،*RESEARCH ARTICLE-COMPUTER ENGINEERING AND COMPUTER SCIENCE* ،Medina .2021 ،

[13] Y .ÇELİK" ،SENTIMENT ANALYSIS OF ARABS IN TURKEY USING DEEP LEARNING ON SOCIAL MEDIA DATA "، KARABUK.2022 ،

[14] Farghaly ،A. Shaalan and K" ، Arabic natural language processing: Challenges and solutions "،*ACM Transaction on Asian Language Information Processing* .2010 ،

[15] Heikal ،Maha ، Torki ،Marwan ، El-Makky Nagwa" ،Sentiment Analysis of Arabic Tweets using Deep Learnin "،*The 4th International Conference on Arabic Computational Linguistics (ACLing 2018* ، Egypt .2018 ،

[16] Toan ،D. Trinh ، Truong and Viet-Hung ، "Support Vector Machine for Short-Term Traffic Flow Prediction and Improvement of Its Model Training using Neares Neighbor Approach "،*Transportation research record* ، Hanoi .2020 ،

[17] F. Pedregosa ،G. Varoquaux ،A. Gramfort ، V. Michel and B. Thirion" ،Scikit-learn: Machine Learning in Python "،*Journal of Machine Learning Research* ،Brucher .2011 ،