استخدام أشجار الغابات العشوائية للتنبؤ بحركة أسعار الذهب

غسان ساكت*، ضحى عداس**

*قسم الإحصاء ونظم المعلومات، كلية الاقتصاد، جامعة حلب

**قسم المحاسبة، كلية الاقتصاد، جامعة حلب

الملخص

هدف البحث إلى استخدام طريقة أشجار الغابات العشوائية للتنبؤ بحركة أسعار الذهب ومقارنة نتائجها بنتائج التحليل التمييزي وأشجار القرار. تم الاستفادة من بيانات أسعار الذهب التاريخية وهي سعر الافتتاح وسعر الاغلاق وأقل سعر وأعلى سعر خلال ساعة كمتغيرات مستقلة من أجل حساب المتغير التابع وهو حركة سعر الذهب بعد ساعة، حيث تم وضع تسعة قرارات كقيم للمتغير التابع تساعد متخذ القرار على اتخاذ القرار الاستثماري الصحيح. تم بناء النماذج واختبار نتائج الطرق المختلفة المستخدمة في التنبؤ ومقارنتها مع بعضها البعض. توصل البحث إلى أن استخدام أشجار الغابات العشوائية تعطي دقة مقدارها 82.727 % للتنبؤ بحركة سعر الذهب مقابل الدولار ودقة مقدارها 80.0135 % للتنبؤ بحركة سعر الذهب مقابل اليورو. كما توصل البحث إلى أن نماذج أشجار القرار (C4.5) تعطي نتائج قريبة من نموذج أشجار الغابات العشوائية مع أفضلية لأشجار الغابات العشوائية. في حين فشلت طريقة التحليل التمييزي في إعطاء نتائج موثوقة لعدم استيفائه للشروط الخاصة به.

الكلمات المفتاحية: أسعار الذهب، أشجار الغابات العشوائية، أشجار القرار، التحليل التمييزي.

ورد البحث للمجلة بتاريخ 2019/3/17

قبل للنشر بتاريخ 2019/6/23

Using Random Forest Trees to Predict the Movement of Gold Prices

Ghassan Saket*, Doha Addas**
*Dept. of Statistic and Information Systems, Faculty of Economics, University of Aleppo
**Dept. of Accounting, Faculty of Economics, University of Aleppo

Abstract

The aim of the research is to use the method of random forest trees to predict the movement of gold prices and compare their results with the results of discriminatory analysis and decision trees. The research used the historical gold price data which are the opening price, the closing price, the lowest price and the highest price within an hour, as independent variables for calculating the dependent variable which is the gold price movement after an hour. Nine decisions were placed as values for the dependent variant for helping the decision maker to make the right investment decision. The models are built and the results of the different methods used in forecasting are tested and compared to each other.

The research found that the use of random forest trees gives 82.727% accuracy to predict the gold price movement against the dollar and 80.0135% accuracy to predict the gold price movement against the euro. The research also found that the models of decision trees (C4.5) give results close to the model of the random forest trees, while the discriminant analysis method failed to produce good results

Keywords: discriminant analysis, Decision Trees, Random Forest, Gold Price Movement.

Received 17/3/2019
Accepted 23/6/2019

مقدمة:

يعتبر الذهب من أهم المعادن وأكثرها شيوعاً وفائدةً لكل من الأشخاص والشركات والبنوك والحكومات لما لهذا المعدن من أهمية كبيرة على كافة الأصعدة، فيستطع الأشخاص والشركات تحقيق أرباح عند القدرة على التنبؤ بسعر أو حركة الذهب، كما تستطيع البنوك التحوط من المخاطر ومواجهة المواقف الصعبة باللجوء إلى الذهب، بالإضافة إلى اعتماد الحكومات على الذهب في تحقيق استقرار في سياستها وذلك لاعتماد اقتصاد الدول على ما تمتلكه من احتياطي من معدن الذهب واستخدام هذا المعدن لمواجهة مشكلة التضخم ومشاكل عدم استقرار أسعار العملة المحلية والمشاكل الاقتصادية.

تلجأ الدول المتقدمة إلى زيادة مخزونها من الذهب كوسيلة للتحوط من المخاطر ومواجهة أي مشكلة مالية او اقتصادية، كما تلجأ الدول النامية في حال تحسن اقتصادها إلى زيادة مخزونها من هذا المعدن وبالتالي تحقيق طلب إضافي على هذا المعدن، وفي حال تراجع اقتصادها إلى بيع جزء من مخزونها لمواجهة مشاكلها، مما يؤدي إلى حركة أسعار هذا المعدن صعوداً أو هبوطاً، وبالتالي يكون الطلب مستمر ومتزايد على هذا المعدن، من هنا تنبع أهمية دراسة أسعار هذا المعدن.

وبشكل عام تتأثر حركة أسعار الذهب بمجموعة من المحددات منها ما هو عالمي ومنها ما هو محلي. تتعدد العوامل المحددة لحركة أسعار الذهب، فمنها المشاكل الاقتصادية العالمية، سعر الفائدة، العرض والطلب الدولي، الاستقرار السياسي (الحروب والنزاعات)، اتجاهات أو رغبات المستثمرين، أسعار الدولار، أسعار اليورو، بالإضافة إلى عوامل عديدة أخرى. ومن الطرق التي تساعد على التنبؤ بحركة أسعار الذهب هي دراسة هذه الأسعار خلال فترة زمنية محددة بأسلوب قائم على أسس علمية صحيحة، من هنا انطلقت فكرة البحث في استخدام الأساليب الإحصائية التقليدية والأساليب التي تعتمد على الذكاء الصنعي للتنبؤ بحركة أسعار الذهب والمقارنة بين الطريقتين.

تم من خلال هذا البحث استخدام بيانات عن أسعار الذهب مقابل الدولار خلال فترة زمنية محددة، كما تم تهيئة هذه البيانات لتناسب الطرق الإحصائية التقليدية وطرق الذكاء الصنعي.

1. مشكلة البحث:

يواجه الأفراد أو الشركات أو الحكومات صعوبة كبيرة اثناء عملية بيع أو شراء معدن الذهب وذلك بسبب صعوبة التنبؤ بسعر الذهب لفترات زمنية لاحقة باستخدام طرق التنبؤ التقليدية، بالإضافة إلى وجود عوامل عديدة تؤثر في حركة سعر الذهب يصعب على متخذ القرار الاستثماري معرفتها أو تحديدها أو التنبؤ بها.

2. أهمية البحث:

إن التنبؤ بحركة سعر الذهب ذو أهمية كبيرة ويساعد على تقديم الدعم والمساعدة للمستثمر لاتخاذ القرار الأفضل وبالوقت المناسب، حيث أن القرار الخاطئ المتعلق بقرار الاستثمار في الذهب يؤثر بشكل سلبي على المستثمر سواء أكان المستثمر فرد أو شركة أو حكومة، لذلك يعد استخدام تقنيات وأدوات حديثة تساعد على اتخاذ القرار الاستثماري بشكل صحيح وبالوقت المناسب وتكلفة منخفضة هو أمر بالغ الأهمية.

3. أهداف البحث:

يهدف البحث إلى إيجاد نموذج قادر على دعم قرار المستثمر في معدن الذهب وذلك من خلال التنبؤ بحركة سعر الذهب مقابل كل من عملتي اليورو والدولار بدقة وزمن مناسبين وذلك باستخدام تقنيات أشجار الغابات العشوائية وأشجار القرار بدلاً من الاعتماد على الطرق التقليدية، وبالتالي مساعدة المستثمر في تخفيض حالة عدم التأكد واتخاذ القرار المناسب.

4. منهجية البحث:

يعتمد البحث على المنهج التحليلي، حيث تم تحليل البيانات الخاصة بأسعار الذهب مقابل كل من عملتي الدولار واليورو باستخدام الأداة WEKA الإصدار 3.8 لبناء نماذج للتنبؤ بحركة الأسعار عن طريق التحليل التمييزي

(كأحد الطرق التقليدية المستخدمة في عملية التصنيف)، أشجار القرار المبنية على أساس تقنية C4.5، وأشجار الغابات العشوائية، كما تم التأكد من قدرة وكفاءة ودقة كل من النماذج السابقة وتحديد النموذج الأفضل.

5. عينة البحث

تم جمع بيانات أسعار الذهب مقابل كل من الدولار واليورو في الفترة الممتدة من 2-1-2013 ولغاية 31 -12- 2017 من موقع history data الجزء الخاص بسوق الفوركس حيث تم الحصول على سعر الافتتاح وأعلى سعر وأقل سعر وسعر الإغلاق لفترة زمنية عبارة عن ساعة واحدة [1].

6. الدراسات السابقة:

أولاً: دراسة (سامية خالد حسن، اوات سردار وادي) التنبؤ بأسعار الذهب وذلك باستخدام سلاسل ماركوف(MLE) [2]

هدفت الدراسة إلى دراسة أسعار الذهب في الفترة من 2008 وحتى 2012 وتم تصنيف أسعار الذهب في ثلاثة مجموعات هي ارتفاع وانخفاض واستقرار، وبعد تطبيق سلاسل ماركوف توصلت الدراسة إلى أن أسعار الذهب باتجاه الصعود، لذلك أوصت الدراسة بالاستفادة من هذه المعلومة في اتخاذ القرارات الاستثمارية.

ثانياً: دراسة (أبو عابدة الفت فتحي سالم) استخدام الطرق الإحصائية في التنبؤ بأسعار الذهب العالمية [3]

استخدمت الدراسة البيانات المتاحة من قبل البنك الدولي للفترة الزمنية من 1985 وحتى 2014 وهي بيانات شهرية لأسعار الذهب مؤلفة من 360 سجل، تم استخدام طريقة بوكس جنكينز وطريقة ((ARIMA (p, d, q) الشبكات العصبية الاصطناعية للتنبؤ بأسعار الذهب، وتوصلت الدراسة على أن الشبكات العصبية الاصطناعية أكثر قدرة على التنبؤ بأسعار الذهب من طريقة بوكس جنكينز.

ثالثاً: دراسة (Nwachukwu & Kabari1) نظام دعم قرار باستخدام شجرة القرار والشبكات العصبية [4]

هدفت الدراسة إلى تطوير نموذج شجرة قرار عصبية للقيام بعملية دعم قرار

نظام منح القروض إضافة إلى نظام دعم قرار سريري لتشخيص أمراض العين اعتماداً على شكوى المريض والأعراض والاختبارات الفيزيائية. يُكامل النظام المقترح بين شجرة القرار والشبكات العصبية الاصطناعية متعددة الطبقات وذات التغذية الأمامية وخوارزمية الانتشار الخلفي في التعلم لبناء النموذج. لتطوير النظام المقترح في الدراسة تم استخدام منهجية التحليل والتصميم غرضية التوجه، كما تم استخدام لغة برمجية غرضية التوجه ضمن الأداة MATLAB لتنفيذ نموذج الدراسة المُقترح. بلغت معدل نجاح النموذج 88%.

رابعاً: دراسة (Kanade Vivek, et al) التنبؤ بالأسواق المالية: باستخدام تحليل البيانات التاريخية [5]

استخدمت الدراسة التحليل الأساسي والتحليل الفني، تم استخدام التحليل الأساسي وتطبيقه على بيانات تم الحصول عليها من وسائل الاعلام الاجتماعي والتي أصبح لها تأثر كبير وخاصة في عصرنا الحالي لتساعد على التنبؤ باتجاه سوق الأسهم، في حين تم استخدام البيانات التاريخية للأسعار في التحليل الفني والذي اعتمد على التحليل الألي للبيانات. اعتمدت طريقة البحث على تضمين الأخبار التي تم جمعها بالإضافة إلى بيانات وسائل الاعلام الاجتماعي للتوصل إلى أحساس الفرد باتجاه السوق، ثم تم تحليل العلاقة بين أحساس الفرد وأسعار الأسهم. توصل البحث إلى إمكانية استخدام النموذج الذي تم تدريبه للتنبؤ بقيم مستقبلية لقيم الأسهم. اعتمد البحث على استخدام الشبكات العصبية الاصطناعية وآلة شعاع الدعم (Support Vector Machine)، وأظهرت الدراسة أن آلة شعاع الدعم ذات قدرة أعلى على التنبؤ من الشبكة العصبية الاصطناعية وأيضاً ذات متوسط مربعات أخطاء أقل.

ما يميز البحث عن الدراسات السابقة هو استخدام أسعار الذهب مقابل عملتين أساسيتين هما الدولار واليورو بالإضافة إلى استخدام تسعة قيم تصنيفية لحركة الأسهم لقيمة المتغير التابع (الهدف)، كما تميز البحث باستخدام أشجار الغابات العشوائية وهو ما لم تتعرض له الدراسات السابقة، إضافة إلى تميز البحث

باستخدام طريقة تقليدية للتصنيف (التحليل التمييزي) وطريق ذكية (أشجار القرار واشجار الغابات العشوائية) والمقارنة بين نتائج النماذج المختلفة.

7. أسعار الذهب:

يعتبر الذهب من أهم أنواع المعادن وأكثرها تداولاً لما له من استقرار في ظل التقلبات الاقتصادية، فيما إذا قورن مع الفضة أو البلاتين أو البلاديوم لاعتباره مؤشراً اقتصادياً يساعد على فهم وتفسير الحالات الاستثمارية في حال طرأت تغيرات مفاجئة على أسعاره ومخزوناً احتياطياً من أجل الحفاظ على التوازن الاقتصادي للدول، لذلك أصبح التداول بالذهب عامل جذب للكثير من المستثمرين للحماية من الحالات التي تنخفض فيها قيمة العملات الورقية خلال فترات الضغوط المالية أو عدم الاستقرار السياسي[6] .

يعد سوق الفوركس من أهم الأسواق المالية الذي يتيح المتاجرة بالمعادن النفيسة كالذهب لاعتباره شكلاً من أشكال العملات، وكلمة FOREX تعني البورصة العالمية للعملات الأجنبية، وهي اختصار (Foreign Exchange Market) أي سوق تداول العملات الأجنبية، وهو سوق وحيد عالمي لا مركزي ليس له مكان محدد في العالم من أجل القيام بالعمليات التجارية، بل يقوم المستثمرون (أشخاص– بنوك– شركات) بعمليات التداول بشكل الكتروني عبر شبكة الانترنت من خلال منصة تداول الكترونية خاصة بشركات الوساطة على مبدأ OTC (over-the- counter) [7].

كما يعتبر سوق الفوركس هو السوق الوحيد عالمياً الذي يتم فيه تداول الذهب مقابل كل من عملتي الدولار واليورو، على عكس الأسواق المنتشرة في كل بلد من بلدان العالم والتي يتم فيها تحديد سعر الذهب مقابل العملة المحلية للبلد.

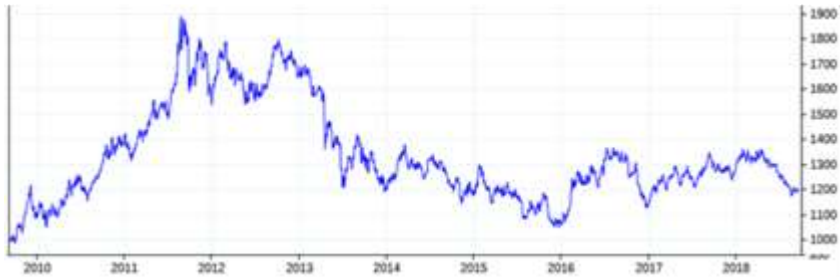
المتاجرة بالذهب في سوق العملات الأجنبية:

تتم المتاجرة في سوق العملات الأجنبية بالأزواج، أي عملة ما مقابل عملة أخرى بالقيمة، وباعتبار الذهب عملة ثمينة وليس مجرد معدن فيتم تداوله مقابل عملات أخرى وكمثال على ذلك الزوج XAU/USD الذهب والدولار الأميركي الأكثر تداولاً في سوق الـ FOREX، إذ تسمى العملة على اليسار XAU بالعملة الأساسية

Base Currency والعملة على اليمين USD بالعملة الثانوية أو العملة المقابلة Counter Currency، فإذا كان سعر صرف الذهب مقابل الدولار الأميركي 1160.9 فإنه يمكننا شراء 1 أونصة (واحدة لقياس كتلة المعادن، تساوي 31.1034768 غرام) من الذهب مقابل كل 1160.9 دولار، أي إن سعر الصرف يمثل المقدار المطلوب دفعة من العملة المقابلة للحصول على وحدة واحدة من عملة الأساس وكذلك بالنسبة لباقي الأزواج بالسوق [8].

هنالك العديد من المزايا التي يتمتع بها سوق العملات الأجنبية، ومن أهمها عملية التداول وتحقيق الأرباح في كلا اتجاهي السوق، أي أن السوق يتيح المتاجرة في كلا الاتجاهين من أجل تحقيق الأرباح سواءً كانت اتجاه حركة الأسعار تصاعدية أو تنازلية، وفي حال كان سعر الصرف باتجاه الصعود هذا يعني أن العملة الأساسية تزداد قوة مقابل العملة الثانوية والعكس صحيح، لأن الدخول في عملية شراء لإحدى العملات هو عملية بيع لعملة أخرى[9] .

الشكل التالي يعبر عن أسعار أونصة الذهب مقابل الدولار خلال الفترة الزمنية من 2010 وحتى نهاية 2018[10] .



الشكل رقم (1) أسعار أونصة الذهب مقابل الدولار من 2010 وحتى 2018

8. أشجار القرار:

تستطيع تقنيات التنقيب في البيانات بناء أشجار قرار ذات كفاءة ودقة عالية باستخدام طرق مختلفة منها C4.5 و ID3 M5P وأشجار الغابات العشوائية وشبكات بييز. وقد تم في هذا البحث استخدام نوعين من أشجار القرار هما C4.5 واشجار الغابات العشوائية.

تبني أشجار القرار هيكل شجري لتمثيل القواعد التي تم استخلاصها من

عملية التصنيف، وتتكون شجرة القرار من مجموعة من العقد تمثل نقط اختبار، وتنتج عنها فروع تمثل مسار، تؤدي في نهايتها الى أوراق تمثل القرار. عقد الاختبار في شجرة القرار هي الحقول التي تعبر عن المتغيرات المستقلة الموجودة في قاعدة البيانات، أما الفروع التي تخرج من هذه العقد فهي عبارة عن القيمة التي تحويها هذه الحقول، في حين تمثل أوراق الشجرة القيم الخاصة بالمتغير التابع (الهدف) [11].

9. أشجار الغابات العشوائية:

قام الباحث Breiman [12] باقتراح اشجار الغابات العشوائية عام 2001 في جامعة كاليفورنيا، وتعتمد بشكل أساسي على خوارزمية CART لبناء أشجار القرار، حيث تبني الغابات العشوائية نتائجها من نتائج أشجار القرار الفردية المستقلة التي يتم بنائها وتعليمها من مجموعات جزئية من بيانات التدريب ثم يتم دمج هذه النماذج، وبالتالي تعد أشجار الغابات العشوائية أداة جديدة وقوية وكثيرة الاستعمال لاستكشاف البيانات وتحليلها ووضع نموذج تنبؤي من خلالها.

يرى Breiman أنه في أشجار الغابات العشوائية لا داعي لاختبار التصديق المتقاطع cross-validation ولا لمجموعة بيانات منفصلة للاختبار والتأكد من خلال قيمة الخطأ التي نحصل عليها من تطبيق النموذج على مجموعة بيانات الاختبار، لأن اشجار الغابات العشوائية تقيّم داخلياً أثناء بنائها، فكل شجرة يتم بناءها من خلال أخذ عينة عشوائية من البيانات لكن يستخدم فقط ثلثي بيانات العينة والثلث الباقي لا يستخدم في بناء الشجرة ويسمى Out-Of-Bag ويوضع أسفل الشجرة لاستخدامه في اختبارها للحصول على تقديرات غير متحيزة ويتم تصنيف البيانات بالصنف الذي يحصل على أعلى نسبة تصويت وهكذا تتكرر العملية حتى يتم بناء الغابة. يؤثر على أداء اشجار الغابات العشوائية إلى حد كبير متغيران رئيسيان فقط هما عدد الأشجار وعدد المتغيرات التي تم اختيارها بشكل عشوائي في كل عملية تقسيم للبيانات.

تتميز الغابات العشوائية بالعديد من المزايا أهمها [13]:

1. تعطي تصنيفات بدقة عالية.

2. تعمل بشكل كفوء على قواعد البيانات الكبيرة.
3. يمكن أن تعالج آلاف المتغيرات المدخلات بدون حذف متغير.
4. تعطي تقييمات حول أهمية المتغيرات في التصنيف.
5. تعطي تقديرات غير متحيزة داخلياً لأنها تعالج خطأ التعميم أثناء بناء الغابة.
6. طريقة فعالة لتقدير البيانات المفقودة وتبقى دقة النموذج عالية حتى عندما تكون نسبة البيانات المفقودة كبيرة.
7. تعطينا طريقة تجريبية من أجل اكتشاف تفاعلات المتغير.
8. نماذج الغابات الناتجة تعطي معلومات حول العلاقة بين المتغيرات والتصنيف.
9. تستطيع غابات الأشجار من خلال استخدام عدة طرائق موازنة الخطأ حتى عندما تكون بيانات الصنف مشوشة في قاعدة البيانات.
10. يمكن حفظ الغابات التي تم بنائها من أجل أن تستخدم لاحقاً للتنبؤ على البيانات الأخرى.
11. حساب التقارب بين السجلات من أجل تجميع الأشجار وتحديد القيم المتطرفة وهذا طريقة مفيدة في عملية تجميع الأشجار لبناء الغابة وتحديد البيانات المتطرفة وإعطاء معلومات مفيدة عن البيانات.

10. تقييم النموذج:

قبل استخدام أي نموذج في عملية التنبؤ يجب التأكد من صحة هذا النموذج، أي يجب تقييمه وعرض نتائج التقييم أمام متخذ القرار. وللتأكد من أن الفروقات التي تم الحصول عليها عن أداء النماذج المختلفة ليست مجرد صدفة، يجب تطبيق أساليب واختبارات إحصائية أخرى. وللقيام بعملية التقييم بشكل صحيح يتم تقسيم البيانات إلى مجموعتين، الأولى للتدريب وبناء النموذج، والثانية لاختبار النموذج.

مقاييس تقييم أداء التصنيف:

من أهم مقاييس تقييم النماذج المستخدمة في عملية التنبؤ والتي طبيعة مخرجاتها عبارة عن أصناف هي نسبة التصنيف الصحيح ونسبة التصنيف الخاطئ

ومعامل Kappa ومعامل ROC Area [14].

تقيس إحصائية Kappa نسبة النجاح المثالية للتصنيف، فهي تقيس التوافق بين تصنيف النموذج والقيم الفعلية بعد تطبيع أو طرح التوافق العشوائي الناتج عن الصدفة. يأخذ هذا المعامل قيماً محصورة بين المجال [0,1]، عندما تكون القيمة قريبة من الواحد 1 يكون التوافق مثالي تام بين تصنيف النموذج والقيم الفعلية، أما عندما تكون القيمة قريبة من الصفر فهذا دليل على أن التوافق عشوائي ناتج من الصدفة وبالتالي لا يمكن الاعتماد على النموذج في عملية التنبؤ. وكلما زادت قيمة هذا المقياس كلما زادت دقة وكفاءة نموذج التصنيف. إذا كانت قيمة Kappa ضمن المجال [0.40–0.59] عندئذ درجة النجاح في التصنيف معتدلة، إذا كانت قيمة Kappa ضمن المجال [0.60-0.79] عندئذ درجة النجاح في التصنيف كبيرة، أما إذا كانت قيمة Kappa ضمن المجال [0.80-1.00] عندئذ درجة النجاح في التصنيف ممتازة [15].

يعبر معامل الـ ROC Area عن قدرة النموذج على تصنيف بيانات الاختبار بشكل صحيح، يأخذ هذا المعامل قيماً محصورة بين المجال [0,1]، ويكون التصنيف أفضل كلما اقتربت قيمته من الواحد، عندما تقترب قيمة هذا المعامل من الـ 0.5 فإنه سوف يكون في منطقة عدم التمييز مما سيؤثر بشكل سلبي على صحة النموذج، أما عندما تكون القيمة قريبة من الصفر فهذا دليل على ضعف النموذج وعدم قدرته على التنبؤ.

11.الدراسة التطبيقية:

12-1 تجهيز وتحضير البيانات:

تم جمع بيانات أسعار الذهب مقابل كل من الدولار واليورو للفترة الممتدة من 2-1-2013 ولغاية 31 -12- 2017 حيث تم الحصول على سعر الافتتاح وأعلى سعر وأقل سعر وسعر الإغلاق لفترة زمنية عبارة عن ساعة واحدة. وقام الباحثان بتهيئة البيانات حيث تم استخدام ثلاث متغيرات للدخل ومتغير للخرج. متغيرات الدخل هي مقدار التغيرات في الأسعار عن الساعة التي تسبقها وهي التغير

في سعر الافتتاح والتغير في أعلى سعر والتغير في أقل سعر، أما متغير الخرج فيدل على التغير في السعر للساعة التي تلي الساعة الحالية. بهذا الشكل أصبح المتغير الخرج يعبر عن الساعة التالية بالاعتماد على قيم متغيرات الساعة الحالية والساعة السابقة لها.

فيما يلي جدول يوضح المتغيرات المستقلة والمتغير التابع التي تم استخدامها في بناء نماذج التنبؤ بحركة أسعار الذهب.

الجدول رقم (1) المتغيرات المستخدمة في البحث- إعداد الباحثان

اسم المتغير	النوع	وصف المتغير
open	مستقل	الفرق بين سعر الافتتاح للساعة الحالية وسعر الساعة السابقة
max	مستقل	الفرق بين أعلى سعر للساعة الحالية وأعلى سعر للساعة السابقة
min	مستقل	الفرق بين أقل سعر للساعة الحالية وأقل سعر للساعة السابقة
move	تابع	الفرق بين سعر الاغلاق للساعة الحالية وسعر الاغلاق للساعة اللاحقة

تم الاعتماد على برنامج الاكسل في تهيئة البيانات، كما تم اسناد 9 قيم للمتغير الخرج (حركة السعر) وهي:

- UpA: في حال التغير في السعر بين الساعة الحالية والساعة التالية من 25 نقطة وما فوق.
- UpB: في حال التغير في السعر بين الساعة الحالية والساعة التالية أقل من 25 وأكبر أو يساوي 15 نقطة.
- UpC: في حال التغير في السعر بين الساعة الحالية والساعة التالية أقل من 15وأكبر أو يساوي 4 نقطة.
- UpD: في حال التغير في السعر بين الساعة الحالية والساعة التالية أقل من 4 نقاط.
- DownA: في حال التغير في السعر بين الساعة الحالية والساعة التالية أقل أو يساوي-25 نقطة.
- DownB: في حال التغير في السعر بين الساعة الحالية والساعة التالية أكبر من −25 وأقل أو يساوي −15 نقطة.

- DownC: في حال التغير في السعر بين الساعة الحالية والساعة التالية أكبر من −15 وأقل أو يساوي −4 نقاط.
- DownD: في حال التغير في السعر بين الساعة الحالية والساعة التالية أكبر من4 وأقل من 0 نقطة.

بعد الانتهاء من تجهيز وتحضير البيانات، تَولدت قاعدتي بيانات واحدة لسعر الذهب مقابل الدولار والأخرى لسعر الذهب مقابل اليورو، كل قاعدة بيانات مؤلفة من 29150 سجل.

تم بناء أكثر من نموذج للتنبؤ من أجل قاعدتي البيانات السابقتين، النموذج الأول باستخدام التحليل التمييزي والثاني باستخدام أشجار القرار والثالث باستخدام أشجار الغابات العشوائية.

عند استخدام كامل بيانات قاعدة البيانات لم تعطي النماذج الثلاث السابقة نتائج يمكن الاعتماد عليها، حيث كان نموذج التحليل التمييزي غير معنوي ولا يمكن الاعتماد عليه، كذلك الأمر بالنسبة لنموذج أشجار القرار(C4.5) واشجار الغابات العشوائية، حيث كانت دقة التصنيف في النماذج الثلاث السابقة بحدود 40 % وهي نسبة قليلة ولا يمكن الاعتماد عليها، وخاصة وأن المؤشرات الإحصائية الخاصة بدقة تصنيف النماذج كانت منخفضة جداً. لذلك قام الباحثان باختيار جزء البيانات الخاص بعام 2017 فقط من كل من قاعدتي البيانات التي تم بنائهما سابقاً. لذلك تم اختيار 5940 سجل من قاعدة البيانات الخاصة بأسعار الذهب مقابل الدولار و 5914 سجل من قاعدة البيانات الخاصة بأسعار الذهب مقابل اليورو. تم استخدام الأداة Weka الإصدار 3.8 لبناء نموذجي أشجار الغابات العشوائية ,أشجار القرار.

12−2 نماذج التنبؤ بأسعار الذهب باستخدام التحليل التمييزي:

قبل تطبيق التحليل التمييزي لابد من التحقق من شروطه والمتمثلة بما يلي [16]:

1. يجب أن يكون حجم العينة كبيراً.
2. عدم وجود ارتباط بين المتغيرات الكمية المستقلة الداخلة في التحليل.
3. المجتمعات الإحصائية محل الدراسة تمثل مجتمعات إحصائية ذات توزيع

طبيعي.

4. أن تكون كل مفردة في المجتمع قابلة للوصف والتحديد بمجموعة من المقاييس أو المتغيرات المستقلة
5. تجانس التباين والتغايير يمكن إهمال هذا الشرط إذا كان حجم العينة أكبر من 100 مشاهدة.
6. عدم تساوي متوسطات المتغيرات المستقلة عند مستويات المتغير التابع.

للتحقق من شروط التحليل التمييزي تم اجراء ما يلي:

تم تشكيل مصفوفة الارتباط بين المتغيرات المستقلة والجدول التالي يبين مصفوفة الارتباط لبيانات الذهب يورو وبيانات الذهب والدولار.

الجدول رقم (2) مصفوفة الارتباط للمتغيرات المستقلة– إعداد الباحثان بالاعتماد على مخرجات spss

		مصفوفة ارتباط بيانات الذهب يورو			مصفوفة ارتباط بيانات الذهب دولار		
		O	H	L	O	H	L
O	Pearson Correlation	1	.971*	.979*	1	.965*	.977*
	Sig. (2-tailed)		.000	.000		.000	.000
H	Pearson Correlation	.971*	1	.961*	.965*	1	.953*
	Sig. (2-tailed)	.000		.000	.000		.000
L	Pearson Correlation	.979*	.961*	1	.977*	.953*	1
	Sig. (2-tailed)	.000	.000		.000	.000	

نلاحظ من مصفوفة الارتباط السابقة أنه يوجد ارتباط قوي بين المتغيرات المستقلة بالنسبة لبيانات الذهب يورو وبيانات الذهب دولار وبالتالي لم يتحقق شرط عدم الارتباط بين المتغيرات المستقلة.

تم إجراء اختبار Kolmogorov-Smirnov للتأكد من أن المتغيرات تتبع التوزيع الطبيعي بالنسبة لبيانات الذهب يورو وبيانات الذهب دولار وتبين أن sig<0.05 لجميع المتغيرات بالنسبة لبيانات الذهب يورو وبيانات الذهب دولار وبالتالي لا تتبع المتغيرات التوزيع الطبيعي ولم يتحقق الشرط الثاني من شروط تطبيق التحليل التمييزي.

تم تطبيق الاختبار التمييزي على بيانات الذهب يورو وبيانات الذهب دولار للتأكد من باقي الشروط وتم اختيار اختبار Box's M للتأكد من تجانس التباين والتغاير وحصلنا على الجدول التالي:

الجدول رقم (3) نتائج اختبار التجانس– إعداد الباحثان بالاعتماد على مخرجات spss

		بيانات الذهب يورو	بيانات الذهب دولار
Box's M		5729.325	4857.367
	Sig.	.000	.000

نلاحظ من اختبار Box's M أن Sig=0.000<0.05 وهذا يدل على عدم تجانس التباين بالنسبة لبيانات الذهب يورو وبيانات الذهب دولار والتغاير ولم يتحقق الشرط الثالث من شروط تطبيق التحليل التمييزي، مع العلم بأنه يمكن التغاضي عن هذا الشرط لأن حجم البيانات أكثر من 100 سجل.

نستنتج مما سبق عدم إمكانية تطبيق التحليل التمييزي على بيانات الذهب يورو وبيانات الذهب دولار لعدم تحقق جميع الشروط الخاصة بتطبيقه.

أولاً– نموذج التنبؤ بحركة أسعار الذهب مقابل الدولار:

الجدول التالي يوضح نتيجة تطبيق اختبار التمييزي على بيانات أسعار الذهب مقابل الدولار رغم عدم تحقق جميع شروط التحليل التمييزي.

الجدول رقم (4) ملخص نتائج تطبيق التحليل التمييزي على بيانات ذهب – دولار– إعداد الباحثان بالاعتماد على مخرجات spss

Kappaمعامل	نسبة التصنيف الصحيح%	Sig اختبار 'Wilks Lambd	اختبار 'Wilks Lambda	الارتباط القانوني	الجذر الكامن
.637	77.9%	.000	.466	.730	1.142a

نلاحظ من الجدول السابق أن قيمة الجذر الكامن تساوي 1.142 وقيمة الارتباط القانوني تساوي 0.730 وبتربيع الارتباط القانوني نحصل على معامل التحديد ويساوي 0.533 وهو مؤشر على نسبة التباين المفسر من قبل التحليل التمييزي، ووفق اختبار Wilks' Lambda للتحقق من معنوية النموذج نلاحظ أن Sig=0.000<0.05 وهذا يدل على أن النموذج معنوي، ونسبة التباين غير المفسر تساوي 0.466، ونسبة التصنيف الصحيح بلغت 77.9%.

وللتأكد من فعالية التنبؤ قمنا بدراسة جدول الاقتران بين المتغير الهدف والمتغير الذي تم التنبؤ به من خلال تحليل التمييزي للحصول على معامل Kappa وقد وصلت قيمتها إلى 0.637.

ثانياً– نموذج التنبؤ بحركة أسعار الذهب مقابل اليورو:

الجدول التالي يوضح نتيجة تطبيق اختبار التمييزي على بيانات أسعار الذهب مقابل اليورو رغم عدم تحقق جميع شروط التحليل التمييزي.

الجدول رقم (5) ملخص نتائج تطبيق التحليل التمييزي على بيانات ذهب – يورو– إعداد الباحثان بالاعتماد على مخرجات spss

Kappaمعامل	نسبة التصنيف الصحيح%	Sig اختبار Wilks' Lambd	اختبار Wilks' Lambda	الارتباط القانوني	الجذر الكامن
.681	81.9%	.000	.478	22.7	1.088a

نلاحظ من الجدول السابق أن قيمة الجذر الكامن تساوي 1.088 وقيمة الارتباط القانوني تسوي 0.722 وبتربيع الارتباط القانوني نحصل على معامل التحديد ويساوي 0.522 وهو مؤشر على نسبة التباين المفسر من قبل التحليل التمييزي، ووفق اختبار Wilks'Lambda للتحقق من معنوية النموذج نلاحظ أن Sig=0.000<0.0 وهذا يدل على أن النموذج معنوي، ونسبة التباين غير المفسر تساوي 0.478، ونسبة التصنيف الصحيح بلغت 81.9%.

وللتأكد من فعالية التنبؤ قمنا بدراسة جدول الاقتران بين المتغير الهدف والمتغير الذي تم التنبؤ به من خلال تحليل التمييزي للحصول على معامل Kappa وقد وصلت قيمتها إلى 0.637.

نلاحظ في كلا النموذجين الخاصين بسعر الذهب مقابل الدولار واليورو بأن نسبة التباين الغير مفسر قريبة من 0.5 وبالتالي من غير المناسب اعتماد نتائج التصنيف.

12–3 نماذج التنبؤ بحركة أسعار الذهب باستخدام أشجار القرار(C4.5):

أولاً– نموذج التنبؤ بحركة أسعار الذهب مقابل الدولار:

تم استخدام بيانات أسعار الذهب مقابل الدولار خلال عام 2017 والمؤلفة من 5940 سجل، كل سجل معبر عن السعر خلال ساعة محددة. وبعد بناء شجرة القرار بالاعتماد على طريقة الاختبار Cross Validation 10حصلنا على النتائج

الموضحة بالجدول التالي:

الجدول رقم (6) نتائج نموذج التنبؤ باستخدام أشجار القرار ذهب – دولار– إعداد الباحثان بالاعتماد على مخرجات Weka

معيار التقييم	قيمة المعامل
نسبة التصنيف الصحيح	82.1717
نسبة التصنيف الخاطئ	17.8283
معاملKapa	0.7252
معاملROC	0.913

بدراسة الجدول السابق نستنتج أن النموذج جيد جداً حيث بلغت نسبة التصنيف الصحيح 82.1717 وهي نسبة جيدة جداً كما بلغ قيمة معامل Kappa 0.7252 وهي قيمة جيدة جداً أيضاً وتدل على أن التصنيف الصحيح لم يكن نتيجة الصدفة، كما أن قيمة معامل الـ ROC مرتفعة وقريبة إلى الواحد وهذا دليل على أن النموذج قادر على التمييز ما بين الأصناف المختلفة لتابع الهدف.

ثانياً– نموذج التنبؤ بحركة أسعار الذهب مقابل اليورو:

تم استخدام بيانات أسعار الذهب مقابل الدولار خلال عام 2017 والمؤلفة من 5914 سجل، كل سجل معبر عن السعر خلال ساعة محددة. وبعد بناء شجرة القرار بالاعتماد على طريقة الاختبار Cross Validation 10 حصلنا على النتائج الموضحة بالجدول التالي:

الجدول رقم (7) نتائج نموذج التنبؤ باستخدام أشجار القرار ذهب – يورو– إعداد الباحثان بالاعتماد على مخرجات Weka

معيار التقييم	قيمة المعامل
نسبة التصنيف الصحيح	79.573
نسبة التصنيف الخاطئ	20.426
معاملKapa	0.7054
معاملROC	0.913

بدراسة الجدول السابق نستنتج أن النموذج جيد حيث بلغت نسبة التصنيف الصحيح 79.573 وهي نسبة جيدة كما بلغ قيمة معامل Kappa 0.7054 وهي

قيمة جيدة جداً أيضاً وتدل على أن التصنيف الصحيح لم يكن نتيجة الصدفة، كما أن قيمة معامل الـ ROC مرتفعة وقريبة إلى الواحد وهذا دليل على أن النموذج قادر على التمييز ما بين الأصناف المختلفة لتابع الهدف.

3-12 نماذج التنبؤ بحركة أسعار الذهب باستخدام أشجار الغابات العشوائية:

أولاً- نموذج التنبؤ بحركة أسعار الذهب مقابل الدولار

تم استخدام بيانات أسعار الذهب مقابل اليورو خلال عام 2017 والمؤلفة من 5940 سجل، كل سجل معبر عن السعر خلال ساعة محددة. وبعد بناء أشجار الغابات العشوائية بالاعتماد على طريقة الاختبار Cross Validation 10 حصلنا على النتائج الموضحة بالجدول التالي:

الجدول رقم (8) نتائج نموذج التنبؤ باستخدام أشجار الغابات العشوائية ذهب – دولار- المصدر: من إعداد الباحثان بالاعتماد على مخرجات Weka

معيار التقييم	قيمة المعامل
نسبة التصنيف الصحيح	82.273
نسبة التصنيف الخاطئ	17.727
معاملKapa	0.727
معاملROC	0.972

بدراسة الجدول السابق نستنتج أن النموذج جيد جداً حيث بلغت نسبة التصنيف الصحيح 82.273 وهي نسبة جيدة جداً كما بلغ قيمة معامل Kappa 0.727 وهي قيمة جيدة جداً أيضاً وتدل على أن التصنيف الصحيح لم يكن نتيجة الصدفة، كما أن قيمة معامل الـ ROC مرتفعة وقريبة إلى الواحد وهذا دليل على أن النموذج قادر على التمييز ما بين الأصناف المختلفة للتابع الهدف.

ثانياً- نموذج التنبؤ بحركة أسعار الذهب مقابل اليورو:

تم استخدام بيانات أسعار الذهب مقابل الدولار خلال عام 2017 والمؤلفة من 5914 سجل، كل سجل معبر عن السعر خلال ساعة محددة. وبعد أشجار الغابات العشوائية بالاعتماد على طريقة الاختبار Cross Validation 10 حصلنا على النتائج الموضحة بالجدول التالي:

الجدول رقم (9) نتائج نموذج التنبؤ باستخدام أشجار الغابات العشوائية ذهب – يورو- المصدر: من إعداد الباحثان بالاعتماد على مخرجات Weka

معيار التقييم	قيمة المعامل
نسبة التصنيف الصحيح	80.0135
نسبة التصنيف الخاطئ	19.9865
معاملKapa	0.7054
معاملROC	0.971

بدراسة الجدول السابق نستنتج أن النموذج جيد حيث بلغة نسبة التصنيف الصحيح 80.0135 وهي نسبة جيدة جيداً كما بلغ قيمة معامل Kappa 0.7054 وهي قيمة جيدة جداً أيضاً وتدل على أن التصنيف الصحيح لم يكن نتيجة الصدفة، كما أن قيمة معامل الـ ROC مرتفعة وقريبة إلى الواحد وهذا دليل على أن نموذج أشجار الغايات العشوائية قادر على التمييز ما بين الأصناف المختلفة للتابع الهدف.

كما تم الملاحظة من مخرجات برنامج الـ weka أنه في جميع النماذج السابقة تجاوزت دقة التنبؤ بكل من الصنفين UpA و DownA الـ 92% في حين أخفقت جميع النماذج في تصنيف باقي قيم التابع الهدف. من هنا نستنتج بأن جميع النماذج قادرة على دعم القرار عندما يكون هناك فرق يمكن أن يستفيد منه المستثمر (أكثر من 25 نقطة ربح أو 25 نقطة خسارة) ويحقق منه ربح أو يتخلص من خسارة. في حين أنه من غير المهم جداً للمستثمر التنبؤ بحالة أقل من 25 نقطة ربح أو 25 نقطة خسارة.

12. الاستنتاجات:

1. عدم قدرة نموذج التحليل التمييزي على التنبؤ بحركة أسعار الذهب مقابل كل من عملتي الدولار واليورو لعدم استيفاء جميع الشروط إضافة النسبة المرتفعة للتابين الغير مفسر.
2. قدرة أشجار الغابات العشوائية واشجار القرار (C4.5) على التنبؤ بحركة اسعار الذهب وبدرجة تأكد جيدة جداً من أجل كل من عملتي الدولار واليورو.
3. عدم وجود فروقات كبيرة بين نماذج اشجار الغابات العشوائية وأشجار القرار على

التنبؤ بحركة أسعار الذهب مع أفضلية لأشجار الغابات العشوائية.

13. التوصيات:

1. نوصي الأفراد والشركات والمستثمرين استخدام نماذج أشجار الغابات العشوائية للتنبؤ بحركة أسعار الذهب مقابل كل من عملتي الدولار واليورو.
2. استخدام تقنيات الذكاء الصنعي للتنبؤ بسعر الذهب بدلاً من حركة سعر الذهب وذلك باستخدام تقنيات أشجار القرار M5p والشبكات العصبية الاصطناعية وذلك للحصول على نتائج أكثر دقة.

المراجع:

[1] https://histdata.com/download-free-forex-data/

[2] حسن سامية خالد، اوات سردار وادي 2013، **التنبؤ بأسعار الذهب وذلك باستخدام سلاسل ماركوف(MLE)**، العدد الرابع، مجلة الدنانير، الجامعة العراق.

[3] أبو عابدة الفت فتحي سالم، 2015 **استخدام الطرق الإحصائية في التنبؤ بأسعار الذهب العالمية**، رسالة ماجستير، قسم الإحصاء، كلية الاقتصاد والعلوم الإدارية، جامعة الأزهر، غزة.

[4] Kabari1 G., Nwachukwu E., 2013, **Decision Support System Using Decision Tree and Neural Networks**, Computer Engineering and Intelligent Systems, Vol.4, No.7.

[5] Kanade Vivek, et al, 2017, **Stock Market Prediction: Using Historical Data Analysis**, International Journal of Advanced Research in Computer Science and Software Engineering, Volume 7, Issue 1

[6] https://fbs.ae/blog/39-تداول-الذهب-في-الفوركس. (19/01/2018)

[7] https://ar.wikipedia.org/wiki/سوق_صرف_العملات (20/02/2018)

[8] Dicks J., 2010, **Forex Trading Secrets**, by McGraw-Hill Companies, P3.

[9] Norris J., et al., 2010, **Mastering the Currency Market**, the McGraw-Hill Companies, Inc, P9-10.

[10]https://investinghaven.com/commodities-gold/gold-price-forecast-2019/

[11] Mayatt G., Johnson W., 2014 - Making Sense of Data I. A Practical Guide to Exploratory Data Analysis and Data Mining.

2nd Edition, Wiley, USA.

[12] Jehad Ali, et al, 2012, "**Random Forests and Decision Trees**", International Journal of Computer Science Issues (IJCSI), Vol. 9, Issue. 5, No 3, pp. 272-278.

[13] Venkatesan .N, Priya .G, 2015, "**A Study of Random Forest Algorithm with implemetation using Weka**" International Journal of Innovative Research in Computer Science and Engineering (IJIRCSE) Vol.1, Issue. 6. pp. 156-162.

[14] Witten I., et al., 2011, **Data Mining Practical Machine Learning Tools and Techniques**, 3rd ed, Morgan Kaufmann

[15] Landis J.; Koch, G., 1977, **The measurement of observer agreement for categorical data**. Biometrics 33: PP. 159–174.

[16] Field. A, "**Discovering Statistics Using SPSS**", 2nd, Sage published, London.