

استخدام تقنيات التعلم الآلي لتحليل البيانات الاقتصادية في سورية لدعم القرارات الاقتصادية

الدكتور خضر العكاري *

الدكتورة بشرى علي **

(تاريخ الإيداع ٢٠٢٤ /٩/١٧ - تاريخ النشر ٢٠٢٥ /١/١٤)

□ ملخص □

يستكشف هذا البحث إمكانية تطبيق تقنيات التعلم الآلي وتحليل البيانات لدعم القرارات الاقتصادية في سورية، وخاصة في ظل التحديات والاضطرابات الاقتصادية التي تواجهها. يسلط البحث الضوء على القصور في النماذج الاقتصادية التقليدية التي تعتمد على منهج الاحتمال الأقصى، ويبرز قدرة التعلم الآلي على التعامل مع البيانات الضخمة وتوضيح أنماط معقدة، وبالتالي تقديم رؤى أعمق وأكثر دقة حول المؤشرات والاتجاهات الاقتصادية. يتناول البحث ثلاث تقنيات رئيسة للتعلم الآلي وتطبيقاتها على البيانات الاقتصادية السورية بالاعتماد على لغتي برمجة R وبايثون: أولاً: يوضح البحث كيفية استخدام الشبكات العصبية، بما في ذلك نماذج الانحدار الذاتي للشبكات العصبية (NNAR) ووحدات البوابة المتكررة (GRU)، للتنبؤ بالمتغيرات الاقتصادية (معدل التضخم ومؤشر سوق دمشق للأوراق المالية). ثانياً: يوضح البحث استخدام الغابات العشوائية لتحليل البيانات النصية المستخرجة من Google Trends، بهدف التنبؤ باتخاذ القرار الاستثماري في سوق دمشق للأوراق المالية بناءً على أنماط البحث عن الكلمات المفتاحية المتعلقة بالاقتصاد السوري. ثالثاً: يتناول البحث مشكلة التنبؤ في بيئة بيانات عالية الأبعاد، ويقترح استخدام انحدار Lasso مع فرض قيود L1 لتقليل المعاملات واختيار المتغيرات الأكثر تأثيراً على مؤشر أسعار المستهلك (CPI)، مما يزيد من دقة التنبؤ ويقلل من التباين. خلص البحث إلى أن تقنيات التعلم الآلي تُعد أدوات فعالة لتحليل البيانات الاقتصادية السورية المعقدة، وتوفر رؤى قيمة لدعم القرارات الاقتصادية. ويوصي البحث بزيادة دمج هذه التقنيات في عمليات التخطيط والتحليل الاقتصادي في سورية لتطوير نماذج تنبؤية أكثر دقة، مما يساهم في اتخاذ قرارات بالسياسة الاقتصادية أكثر استنارة وفعالية، ويساعد على تخفيف حالة عدم اليقين الاقتصادي. **الكلمات المفتاحية:** التعلم الآلي - تحليل البيانات - الشبكات العصبية - الغابة العشوائية - الأبعاد العالية.

* دكتوراه في الإحصاء والبرمجة - قسم الإحصاء والبرمجة - كلية الاقتصاد - جامعة تشرين طرطوس - سورية -

khderalakkari1990@gmail.com

** مدرس - قسم العلوم المالية والمصرفية - كلية الاقتصاد - جامعة طرطوس - طرطوس - سورية - BushraAli@tartous-univ.edu.sy

Using Machine Learning Techniques to Analyze Economic Data in Syria to Support Economic Decisions

Dr. Khder Al-Akkari*

Dr. Bushra Ali**

(Received 17/9/2024.Accepted 14/1/2025)

□ABSTRACT □

This research explores the potential of applying machine learning and data analysis techniques to support economic decision-making in Syria, particularly in light of the economic challenges and disruptions facing the country. The research highlights the limitations of traditional economic models based on the Maximum Likelihood Estimation (MLE) method and emphasizes the ability of machine learning to handle big data and capture complex patterns, thereby providing deeper and more accurate insights into economic indicators and trends.

The research addresses three main machine learning techniques and their applications to Syrian economic data: First, it demonstrates how neural networks, including Neural Network Autoregression (NNAR) models and Gated Recurrent Unit (GRU), can be used to forecast economic variables such as inflation rates and the Damascus Securities Exchange Index. Second, it illustrates the use of random forests to analyze textual data extracted from Google Trends, aiming to predict investment decisions in the Damascus Securities Exchange based on search patterns for keywords related to the Syrian economy. Third, the research tackles the challenge of prediction in a high-dimensional data environment and proposes the use of Lasso regression with L1 penalty to shrink coefficients and select the most influential variables on the Consumer Price Index (CPI), enhancing prediction accuracy and reducing variance. The research concludes that machine learning techniques are effective tools for analyzing complex Syrian economic data and provide valuable insights for supporting economic decision-making. It recommends further integration of these techniques into economic planning and analysis processes in Syria to develop more accurate predictive models, contributing to more informed and effective policy decisions and helping to mitigate economic uncertainty.

Key word: Machine Learning, Data Analysis, Neural Networks, Random Forest, High-Dimensionality.

*PhD in Statistics and Programming - Department of Statistics and Programming - Faculty of Economics - Tishreen University – Tartous, Syria. khderalakkari1990@gmail.com

**Assistant Professor, Department of Financial and Banking Sciences - Faculty of Economics - Tartous University – Tartous, Syria. BushraAli@tartous-univ.edu.sy.

1- المقدمة:

إن إدخال تعلم الآلة في التحليل الاقتصادي في سورية ضرورة في هذا الوقت حيث واجه اقتصاد البلاد في سورية تحديات واضطرابات كبيرة، مما جعل التنبؤ والتحليل الاقتصادي التقليدي أكثر تعقيداً. من خلال الاستفادة من القدرات التنبؤية لتعلم الآلة وقدرتها على التعامل مع كميات كبيرة من البيانات، يسعى هذا البحث إلى تقديم منظور جديد حول المؤشرات والاتجاهات الاقتصادية في سورية خلال الفترة 2010-2023. إن الفوائد المحتملة لتطبيق تعلم الآلة على البيانات الاقتصادية السورية متعددة. ويمكن أن يساعد في تحديد المحركات الرئيسية للنشاط الاقتصادي، والتنبؤ بالسيناريوهات الاقتصادية المستقبلية، وتقييم تأثير قرارات السياسة. علاوة على ذلك، يمكن أن يساعد تعلم الآلة في اكتشاف الحالات الشاذة والمخالفات في البيانات الاقتصادية، وهو أمر بالغ الأهمية للسعي إلى تخفيف حالات عدم التأكد الاقتصادي. سوف يتعمق هذا البحث في تقنيات تعلم الآلة المختلفة، تحليل الانحدار، وخوارزميات التصنيف، والتنبؤ، لتحليل البيانات الاقتصادية السورية. ومن المتوقع أن تسهم النتائج في تقديم رؤى لبعض مؤشرات الاقتصاد السوري مع عدم توافر آني للبيانات، مما يوفر معلومات قيمة حول كيفية تطبيق تعلم الآلة في البيئات الاقتصادية المعقدة مثل تلك الموجودة في سورية.

2- مراجعة الدراسات السابقة

هدفت دراسة (مرهج، 2013) إلى بناء شبكة عصبية اصطناعية ذات خصائص ملائمة اعتماداً على خبرة الباحث وقاعدة التجربة والخطأ بهدف التنبؤ بقيم مؤشر سوق دمشق للأوراق المالية واتجاهاتها ليوم التداول التالي. تم تدريب الشبكة العصبية وفق خوارزمية التكاثر الارتدادي على عدة متغيرات. امتدت فترة الدراسة خلال 2010-2013 والتي تمت معايرتها لتسهيل عملية التدريب. أظهرت النتائج أن الشبكة العصبية المستخدمة تمكنت من التنبؤ بقيم مؤشر سوق دمشق للأوراق المالية ليوم التداول التالي بدرجة عالية من الدقة ودرجة خطأ 1%.

قدمت دراسة (الحמיד وآخرون، 2020) مقترح لبناء نموذج هجين بين الشبكات العصبونية الاصطناعية ونماذج ARIMA-ARCH لتحسين التنبؤ بحركة مؤشر سوق دمشق. تم تقسيم الفترة الزمنية إلى فترتين من الشهر الثامن إلى الشهر التاسع لعام 2019 الثانية للشهر العاشر 2019. توصلت الدراسة إلى أن الأسلوب المقترح كان الأنسب للتنبؤ باتجاه حركة مؤشر سوق دمشق، إذ لا يعتمد هذا النموذج في حسابه لأوزان الدمج على أخطاء النماذج في مرحلة التقدير وإنما على قدرة النماذج على محاكاة اتجاه الحركة لسلسلة مؤشر السوق وهو ما أعطاه قدرة أكبر في التنبؤ.

هدفت دراسة (يوسف، 2024) إلى التنبؤ بالأرقام القياسية لأسعار المستهلك للفترة القادمة استناداً لبيانات الرقم القياسي لأسعار المستهلك وربطه بسعر صرف الدولار الأمريكي خلال الفترة 2011-2019. تم الاعتماد على منهجية ARDL ونموذج تصحيح الخطأ للتقدير والتنبؤ. أظهرت نتائج الدراسة أن ارتفاع سعر الصرف سيؤدي إلى تضخم مستورد أكبر، وتؤدي التوقعات دوراً مهماً في تفسير الارتفاع المستمر في الأسعار، كما أنه يمكن استخدام نموذج قادر على التنبؤ بشكل كبير حتى فترة ثلاثة أشهر.

ارتكزت دراسة (Khan et al., 2022) إلى تقييم فاعلية تقنيات التعلم الآلي في التنبؤ بالمتغيرات الاقتصادية، وتحديد معدل التضخم وأسعار الصرف في باكستان. اعتمدت الدراسة على منهجية كمية، باستخدام بيانات شهرية للفترة من كانون الثاني 1989 إلى كانون الأول 2020، وقد تم جمع البيانات من مصادر موثوقة مثل

"Thomson Reuters" و "Pakistan Bureau of Statistics". ولتحقيق أهداف الدراسة، تم استخدام أربع خوارزميات تعلم آلي رئيسية، وهي: "K-Nearest Neighbor" (KNN)، و"Polynomial Regression"، والشبكات العصبية الاصطناعية (ANNs)، وآلات المتجهات الداعمة (SVM). تم تقسيم البيانات إلى مجموعتين، بيانات التدريب (1989-2018) وبيانات الاختبار (2019-2020)، وتم تقييم أداء النماذج باستخدام مقياس الجذر التربيعي لمتوسط الخطأ (RMSE) ومتوسط الخطأ المطلق (MAE). أشارت النتائج إلى أن نماذج ANN قدمت أداءً متميزاً في التنبؤ بمعدل التضخم، بينما تفوق نموذج SVM مع دالة "RBF" في التنبؤ بأسعار الصرف. أوصت الدراسة بضرورة تبني تقنيات التعلم الآلي في تحليل البيانات الاقتصادية، لما لها من قدرة على تحسين دقة التنبؤ.

تُشير الدراسات السابقة، مثل دراسة (مرهج، 2013) و(الحميد وآخرون، 2021)، إلى فعالية الشبكات العصبية في التنبؤ بمؤشر سوق دمشق للأوراق المالية. إلا أن هذه الدراسات اقتصرت على نماذج وتقنيات محددة، ولم تتناول التحديات التي تفرضها البيانات عالية التردد أو فترات عدم التأكد الاقتصادي. بالمقارنة، توسع هذا البحث باستخدام نماذج متقدمة مثل GRU، وتقنيات التحقق المتقاطع للسلاسل الزمنية، وتضمن فترات عدم التأكد في التنبؤات، مما يوفر رؤية أكثر دقة وشمولية. أما دراسة (يوسف، 2024)، فتركز على نمذجة العلاقة بين سعر الصرف والتضخم باستخدام منهجية ARDL، أما دراسة (Khan et al., 2022) تعتمد على التنبؤ بترددات شهرية لمتغيرات سعر الصرف والتضخم دون التطرق للترددات المختلفة في الاقتصاد، بينما يعتمد هذا البحث على مجموعة متنوعة من تقنيات تعلم الآلة، مثل الشبكات العصبية، والغابات العشوائية، وانحدار LASSO، لنمذجة وتحليل متغيرات اقتصادية متعددة وذلك بما يتوافق مع خصائص كل بيانات وتكرارها والهدف النهائي ضمن التصنيف أو التنبؤ أو الأثر، بما في ذلك التضخم، ومؤشر سوق دمشق، وعوامل اتخاذ القرار الاستثماري. إضافةً إلى ذلك، يُقدم هذا البحث تحليلاً مفصلاً لتأثير المتغيرات المختلفة على مؤشر أسعار المستهلك باستخدام انحدار LASSO، مع تحديد المتغيرات الأكثر تأثيراً وتقييم دقة النموذج باستخدام تقنيات التحقق المتقاطع. وبالتالي، يُقدم هذا البحث إضافة نوعية للدراسات السابقة من خلال استخدام تقنيات تعلم آلة متقدمة، وتضمن فترات عدم التأكد، وتحليل البيانات النصية، والتعامل مع التحديات التي تفرضها البيانات عالية الأبعاد، مما يُسهم في فهم أعمق للاقتصاد السوري ودعم عملية اتخاذ القرارات الاقتصادية.

3- مشكلة البحث:

بالنظر إلى التحديات الاقتصادية التي تواجهها سورية، والتي تتسم بالتعقيد وعدم اليقين، تبرز الحاجة الملحة إلى تطوير أدوات تحليلية أكثر فعالية وقدرة على استيعاب البيانات الضخمة والمتنوعة والتي تصدر وفق تكرارات وتأخر زمني مختلف، والتي تعجز النماذج الاقتصادية التقليدية القائمة على منهج الاحتمال الأقصى عن التعامل معها بكفاءة. فمن هنا، يبرز السؤال الرئيس الذي يتطلب إجابة: كيف يمكن الاستفادة من تقنيات التعلم الآلي المتقدمة لتحليل البيانات الاقتصادية في سورية، بما يُسهم في تطوير نماذج تنبؤية دقيقة، ويُعزز من فاعلية القرارات الاقتصادية في ظل هذه الظروف؟

4- أهمية البحث وهدفه:

تبرز الأهمية التطبيقية لهذا البحث في قدرته على تقديم أدوات عملية لصناع القرار الاقتصادي في سورية، حيث يُمكن من خلال تطبيق تقنيات التعلم الآلي تحليل البيانات الاقتصادية المعقدة واستخلاص رؤى أعمق حول المؤشرات والاتجاهات الاقتصادية. يتيح هذا التحليل التنبؤ الأكثر دقة بالتطورات الاقتصادية، مما يُمكن من اتخاذ قرارات مستنيرة وفعالة في السياسات الاقتصادية، ويسهم في تخفيف حدة عدم اليقين الاقتصادي. أما على الصعيد النظري، فيسهم هذا البحث في إثراء الأدبيات العلمية من خلال تقييم فاعلية تقنيات التعلم الآلي في سياق اقتصادي خاص، كالاقتصاد السوري، الذي يشهد تحديات واضطرابات كبيرة. كما يقدم البحث نموذجاً لتطبيق هذه التقنيات في بيانات معقدة وعالية الأبعاد، ويُسلط الضوء على أوجه القصور في النماذج الاقتصادية التقليدية، مما يفتح آفاقاً جديدة للبحث والتطوير في هذا المجال. وبالتالي تشمل أهداف البحث:

- ✓ تحديد مدى فاعلية تقنيات التعلم الآلي، كالشبكات العصبية والغابات العشوائية وانحدار Lasso، في تحليل البيانات الاقتصادية السورية والتنبؤ بمتغيراتها الرئيسية، كمعدل التضخم ومؤشر سوق دمشق للأوراق المالية.
- ✓ تطوير نماذج تنبؤية أكثر دقة للبيانات الاقتصادية السورية، من خلال دمج تقنيات التعلم الآلي المتقدمة وتحسين قدرتها على التعامل مع البيانات ذات الأبعاد العالية والتعقيد.
- ✓ تقديم رؤى وتوصيات عملية لصناع القرار الاقتصادي في سورية، من خلال تحليل البيانات الاقتصادية باستخدام تقنيات التعلم الآلي، وتوفير معلومات دقيقة وموثوقة تدعم اتخاذ قرارات اقتصادية أكثر استنارة وفعالية.

5-فرضيات البحث:

- تسهم تقنيات التعلم الآلي في تحسين دقة التنبؤ بالمتغيرات الاقتصادية في سورية. وتتفرع الفرضية الرئيسية إلى:
- تُحقق نماذج الشبكات العصبية (بما في ذلك نماذج NNAR و GRU) دقة تنبؤ أعلى لمعدل التضخم ومؤشر سوق دمشق للأوراق المالية في سورية، مقارنة بالنماذج الاقتصادية التقليدية.
- تسهم الغابات العشوائية في استخلاص رؤى مفيدة من البيانات النصية المستخرجة من Google Trends، مما يُحسن من دقة التنبؤ بقرارات الاستثمار في سوق دمشق للأوراق المالية.
- يُمكن لانحدار Lasso مع فرض قيود L1 من تحديد المتغيرات الاقتصادية الأكثر تأثيراً في مؤشر أسعار المستهلك (CPI) في سورية، وبالتالي تحسين دقة التنبؤ وتقليل التباين.

6-الحدود المكانية والزمانية للبحث:

تم تطبيق الدراسة في سورية خلال الفترة الزمنية 2010-2023.

7- متغيرات البحث:

تم اعتماد التردد الشهري لمتغيرات البحث، وتم ترميز هذه المتغيرات كما يُظهر الجدول (1):

جدول (1): متغيرات الدراسة وترميزها والمصدر:

Variable	Time (Monthly)	Measuring unit	Code	Source
مؤشر سوق دمشق للأوراق المالية	2010-2023	Point	DWX	(Damascus Stock Exchange, 2023)

عدد الشركات المتداولة	2010-2023	Company	NTC	(Damascus Stock Exchange, 2023)
القيمة السوقية	2010-2023	Billion Syrian pounds	MV	(Damascus Stock Exchange, 2023)
قيمة التداول	2010-2023	Billion Syrian pounds	TV	(Damascus Stock Exchange, 2023)
المتوسط اليومي لقيمة التداول	2010-2023	Billion Syrian pounds	ADTV	(Damascus Stock Exchange, 2023)
عدد الأوراق المتداولة	2010-2023	Billion Syrian pounds	NST	(Damascus Stock Exchange, 2023)
عدد الصفقات	2010-2023	Thousand deals	NET	(Damascus Stock Exchange, 2023)
عدد الأيام المتداولة	2010-2023	Day	NTD	(Damascus Stock Exchange, 2023)
مؤشر أسعار المستهلك	2010-2023	Point	CPI	(Central Bureau Of Statistics , 2021)
سعر الصرف الرسمي	2010-2023	Bound	EXR	(Central Bank of Syria, 2022)
معدل الفائدة طويل الأجل	2010-2023	Point	INTM	(Central Bank of Syria, 2022)
أسعار الذهب	2010-2023	Bound	Gold	(Craftsmanship Association, 2023)
الذهب - Google Trend	2010-2023	Point	GTG	Google Trend
الأسهم - Google Trend	2010-2023	Point	GTS	Google Trend
التعليم - Google Trend	2010-2023	Point	GTE	Google Trend
الغاز - Google Trend	2010-2023	Point	GTGg	Google Trend
النفط - Google Trend	2010-2023	Point	GTO	Google Trend
العمالة - Google Trend	2010-2023	Point	GTEM	Google Trend
الاستهلاك - Google Trend	2010-2023	Point	GTC	Google Trend
الزراعة - Google Trend	2010-2023	Point	GTA	Google Trend
الصناعة - Google Trend	2010-2023	Point	GTI	Google Trend
الأسعار - Google Trend	2010-2023	Point	GTP	Google Trend
الاستثمار - Google Trend	2010-2023	Point	GTINV	Google Trend
سعر الصرف - Google Trend	2010-2023	Point	GTEX	Google Trend
اتجاهات - Google Trend	2010-2023	Point	GTD	Google Trend
قوانين - Google Trend	2010-2023	Point	GTL	Google Trend

المصدر: إعداد الباحثان بالاعتماد على بيانات المكتب المركزي للإحصاء ومصرف سورية المركزي والبنك الدولي.

8- منهجية البحث:

يتبع البحث منهج التحليل الإحصائي لتحقيق هدف البحث باستخدام لغة البرمجة الإحصائية R وبايثون. وفقاً للمنهجية المتبعة في البحث يتضمن سير العمل الخطوات الآتية في إطار الاقتصاد القياسي: 1- الانحدار الذاتي للشبكات العصبية، 2- نموذج بوابة الوحدة المتكررة، 3- تقنية الغابة العشوائية، 3- التنبؤ عالي الأبعاد ودالة الخسارة LASSO، 4- التحقق المتقاطع للسلاسل الزمنية. ثم تم تقدير النموذج، ثم عرض النتائج والمناقشة.

9- الأساليب القياسية المستخدمة:

تعتمد التوقعات الاقتصادية تقليدياً على النماذج التي يتم تقديرها باستخدام منهج الاحتمال الأقصى (MLE). القيود المرتبطة بـ MLE معروفة جيداً كما هو موضح في (Bradley E, Trevor, 2021) (يمكن القول إن الاحتمال الأقصى كان أكثر تقنيات الرياضيات التطبيقية تأثيراً في القرن العشرين، ولا يزال هو الأسلوب الأول الذي يختاره الإحصائيون. بشكل عام، يوفر الاحتمال الأقصى تقديرات شبه غير متحيزة بتباين قريب للأدنى، ويفعل ذلك بطريقة تلقائية. ومع ذلك، أظهر تقدير الاحتمال الأقصى نفسه أنه أداة غير كافية وغير ناجحة في العديد من التطبيقات في القرن الواحد والعشرين. مرة أخرى بشكل عام، يمكن أن يكون لعدم التحيز ضرورة لا يمكن تجنبها عندما يكون هناك مئات أو آلاف العوامل لتقديرها في نفس الوقت).

في السنوات الأخيرة، اكتسبت تقنيات التعلم الآلي قوة جذب في مختلف المجالات وهو ما يوضحه تقرير الذكاء الاصطناعي الصادر عن جامعة ستانفورد (Wald, 1949)، بما في ذلك الاقتصاد (Mullainathan and Spiess, 2017; Gogas and Papadimitriou, 2021; Paruchuri, 2021) والتعامل مع مجموعات البيانات واسعة النطاق. أحدث التعلم الآلي (ML) ثورة في عملية صنع القرار في مجموعة متنوعة من المجالات من خلال توفير أدوات جديدة للتنبؤ. في جوهره، ينطوي ML على صياغة دالة خسارة أو تكلفة لقواعد التنبؤ. في هذا السياق، تتوقع قاعدة التنبؤ، المشار إليها بـ $f(x_t)$ قيمة متغير الهدف، $y_t + h$ ، في أفق مستقبلي، h ، استناداً إلى المعلومات المتاحة في الوقت t . تقيس دالة الخسارة، $\ell(y_t + h, f(x_t))$ ، الخطأ الذي ينتج عن القيمة المتوقعة مقارنة بالنتيجة الفعلية. الهدف الرئيس هو تقريب قاعدة القرار الأمثل، f^* التي تقلل الخسارة المتوقعة $E[\ell(y_t + h, f(x_t))]$. تعود جذور هذا النهج إلى نظرية القرار (Granger and Pesaran, 2000) وتم اعتماده في التعلم الإحصائي والتنبؤ الاقتصادي (Chen, 2007). على سبيل المثال، عند استخدام دالة الخسارة التكميلية، $\ell(y_t + h, f(x_t)) = (y_t + h - f(x_t))^3$ تتوافق قاعدة القرار المثالية مع الانحدار (غير الخطي)، $f^*(x_t) = E[y_t + h | x_t]$ فيما يتعلق $f(x_t)$ على وجه التحديد، هذا يعني أن قاعدة التنبؤ المثالية هي تلك التي تقلل من الفرق التربيعي بين القيمة المتوقعة والنتيجة الفعلية.

تستند العديد من الأدوات المستخدمة في التعلم الآلي إلى مفاهيم إحصائية قديمة حيث يمكن ربط بعض الأساليب الشائعة في التعلم الآلي بالنماذج الإحصائية التقليدية وفق الآتي: أولاً، التعلم العميق، وهو أحد أشهر الأساليب في التعلم الآلي. التعلم العميق هو عبارة عن مجموعة من الخوارزميات التي تستخدم شبكات عصبية متعددة الطبقات لمعالجة البيانات ذات الأبعاد العالية، مثل الصور والنصوص والصوت. يمكن فهم التعلم العميق على أنه نموذج انحدار غير خطي (Breiman et al., 1984)، حيث يتم تعديل المعلمات (الأوزان والتحيز) لتقليل دالة الخسارة (الخطأ) بين الإخراج المتوقع والفعلي. يمكن تمثيل كل طبقة في الشبكة العصبية على أنها تحويل غير خطي للطبقة السابقة، وبالتالي يمكن اعتبار الشبكة العصبية كلها على أنها تركيب معقد من الدوال غير الخطية. ثانياً، الغابات العشوائية والتعزيز التدريجي boosting، وهما أسلوبان شائعان في التعلم الآلي لحل مشاكل التصنيف والانحدار. هذه الأساليب تستخدم مجموعة من النماذج البسيطة (عادةً ما تكون أشجار القرار) لإنتاج تنبؤ متوسط أو مرجح (James and Stein, 1961).

يمكن فهم هذه الأساليب على أنها جيل جديد من أشجار الانحدار والتصنيف، والتي تستخدم تقنيات مثل التعزيز لتحسين الدقة. يمكن تمثيل كل شجرة على أنها قاعدة انقسام هرمية تستند إلى معايير محددة، وبالتالي يمكن

اعتبار الغاية كلها على أنها تركيب متوازن من القواعد الهرمية. أخيراً، أسلوب الانحدار المُقَيّد *penalized*، وهو أسلوب في التعلم الآلي يستخدم لمعالجة المشاكل العكسية غير الموصوفة بشكل جيد، مثل الانحدار الخطي متعدد المتغيرات. الانحدار المُقَيّد يضيف قيود إلى دالة الخسارة لتقليل حجم المعلمات ومنع التباين الزائد. يمكن تتبع هذا الأسلوب إلى فكرة الانكماش *shrinkage*، أو التنظيم *regularization*، وانحدار الحرف *ridge* (Persio and Honchar, 2017) يمكن تمثيل الانحدار المُقَيّد على أنه نموذج انحدار خطي مع شرط إضافي على المعلمات، وبالتالي يمكن اعتباره تركيباً بسيطاً من الدوال الخطية مع مجموعة من الشروط. تعتمد قواعد القرار المستخدمة للتنبؤ على البيانات المتاحة بشكل كامل. ومع ذلك، هناك مفاضلة بين التحيز (إعطاء مرونة في دالة تقدير البيانات) والتباين بهدف تحسين أداء التنبؤ بالأداء. من أجل تقليل التباين، يؤدي التنظيم وتقليل الأبعاد أيضاً إلى التحيز. أصبح من الممكن استخدام مجموعة متنوعة من الأدوات عالية الأبعاد وغير المعلمية من خلال التعلم الآلي، والتي يمكن استخدامها لتحسين أداء التنبؤ، والتكيف مع مفاضة التحيز والتباين، وتوفير تقديرات تقريبية مرنة ودقيقة لقواعد القرار المثلى. وهذا يعني أن التعلم الآلي يمكن أن يساعدنا في تطوير قواعد اتخاذ قرار أفضل للتنبؤ من خلال الموازنة بشكل مناسب بين التحيز والتباين. في حين أن تطور أساليب تعلم الآلة له تاريخ طويل، إلا أن النجاح الملحوظ الذي تحقق مؤخراً والاعتماد على نطاق واسع يرجع في الغالب إلى التوافر المتزايد للبيانات الجديدة عالية الأبعاد، والقدرة الحسابية الرخيصة، والحزم الإحصائية القابلة للتطوير. ويعتمد الاقتصاديون أيضاً بشكل متزايد على البيانات عالية الأبعاد. مجموعات البيانات مثل البيانات النصية والمصورة، أو مؤشرات Google. ونتيجة لذلك، تكتسب أساليب تعلم الآلة التقدير وأصبحت منتشرة في كل مكان في الاقتصاد والتمويل.

9-1 الانحدار الذاتي للشبكات العصبية:

نموذج NNAR (الانحدار الذاتي للشبكات العصبية) هو أداة إحصائية متقدمة تجمع بين قوة النماذج الانحدارية الذاتية ومرونة ولا خطية الشبكات العصبية. هذا النهج مفيد بشكل خاص للتنبؤ بالسلاسل الزمنية، حيث يكون من الضروري توضيح الأنماط والديناميكيات المعقدة. في هذا الإطار، يتم بيان الأساس النظري، الصياغة الرياضية، والبنية المحددة لنموذج NNAR (10, 1, 6) المستخدم للتنبؤ بمعدل التضخم في سورية حتى عام 2030.

يحدد نموذج NNAR الانحدار الذاتي التقليدي عن طريق دمج عناصر الشبكة العصبية، مما يسمح له بالنقاط العلاقات غير الخطية في البيانات. الفكرة الأساسية هي استخدام الملاحظات السابقة للتنبؤ بالقيم المستقبلية، حيث يتم إدخال الملاحظات السابقة في شبكة عصبية تحدد التنبؤ. ويتمثل من خلال المكونات التالية:

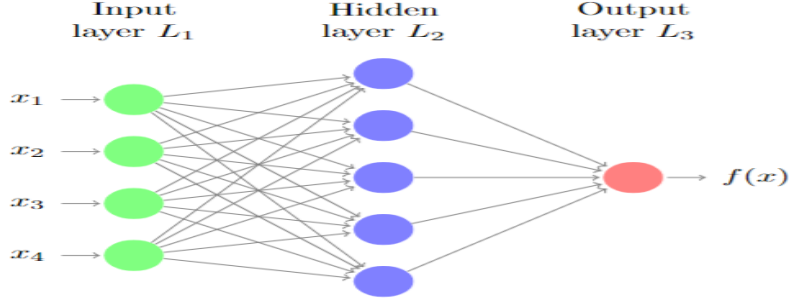
- المكون الانحداري الذاتي: في نموذج AR التقليدي، يتم التعبير عن القيمة المستقبلية للسلسلة الزمنية كمزيج خطي من القيم السابقة. لنموذج $AR(p)$:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \epsilon_t \quad (1)$$

حيث y_t هو القيمة في الزمن t و ϕ هي المعاملات و ϵ_t هو مصطلح الخطأ.

- المكون الشبكي العصبي: يستبدل نموذج NNAR المزيج الخطي بشبكة عصبية تأخذ القيم السابقة كمداخل وتنتج القيمة المستقبلية. هذا يسمح للنموذج بالنقاط العلاقات الأكثر تعقيداً وغير الخطية.

يمكن وصف نموذج NNAR(p,k,m) بشكل رسمي على النحو الآتي: p: عدد المشاهدات المتأخرة (مصطلحات الانحدار الذاتي). k: عدد الطبقات المخفية. m: عدد الخلايا العصبية في كل طبقة مخفية وبوضوح الشكل (1) مكونات الشبكة العصبية.



الشكل (1): شبكة عصبية ذات طبقة مخفية واحدة و 10 خلايا عصبية.

بالنسبة لنموذج NNAR(6,1,10): p = 6: يستخدم النموذج آخر 6 مشاهدات للتعنبؤ بالقيمة التالية. k = 1: يوجد طبقة مخفية واحدة. m = 10: تحتوي الطبقة المخفية على 10 خلايا عصبية. وفق السابق تكون الصيغة العامة لنموذج NNAR (6,1,10) هي:

$$y_t = f(y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}, y_{t-5}, y_{t-6}) + \epsilon_t \quad (2)$$

حيث f هي دالة الشبكة العصبية. ويمكن تفصيل بنية الشبكة العصبية على النحو الآتي:

طبقة الإدخال: تتكون من القيم الست السابقة $(y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}, y_{t-5}, y_{t-6})$ ، **الطبقة المخفية:** تحتوي على 10 خلايا عصبية. كل خلية عصبية تتلقى المدخلات من جميع المشاهدات الست السابقة، يتم معالجتها من خلال دالة تنشيط ReLu وتنتج مخرجاً x إذا كانت x موجبة و 0 بخلاف ذلك، وهي تعطي ميزات من خلال تناثر التنشيط sparsity of the activation مع شبكة عصبية كبيرة بها الكثير من الخلايا العصبية. حيث أنه من الأفضل ألا يتم تنشيط عدد قليل من الخلايا العصبية في الشبكة، وبالتالي جعل عمليات التنشيط منفردة وفعالة. ReLu تعطينا هذه الفائدة. مع شبكة ذات أوزان عشوائية مبدئية (أو طبيعية) وما يقرب من 50% من الشبكة تنتج 0 تنشيط بسبب خاصية ReLu (الإخراج 0 للقيم السالبة ل x). وهذا يعني أن عدداً أقل من الخلايا العصبية يتم إطلاقه (التنشيط المتناثر) وأن الشبكة أخف وزناً. تعطي الدالة من خلال:

$$\sigma(\mu) = \max(0, \mu) \quad (3)$$

رياضياً، يمكن وصف العمليات داخل الشبكة وفق حسابات الطبقة المخفية:

$$h_j = g(\sum_{i=1}^6 w_{ji}y_{t-i} + b_j) \text{ for } j = 1, 2, \dots, 10 \quad (4)$$

حيث g هي دالة التنشيط، w_{ji} هي الأوزان، b_j هي التحيز. ويتمثل الوصف الرياضي وفق حسابات طبقة الإخراج:

$$\hat{y}_t = \sum_{j=1}^{10} v_j h_j + c \quad (5)$$

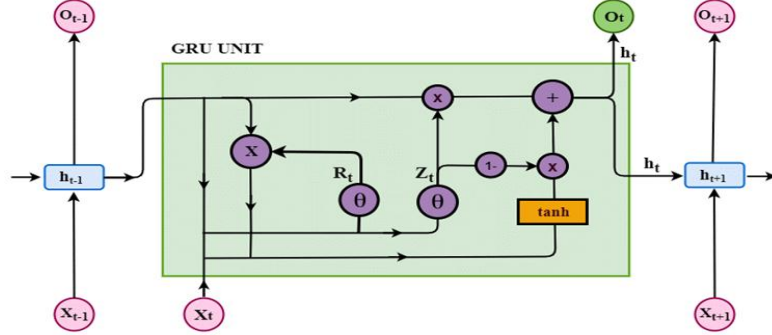
حيث v_j هي الأوزان التي تربط الطبقة المخفية بطبقة الإخراج، و c هو مصطلح التحيز لخلية الإخراج. نموذج NNAR(6,1,10) يجمع بشكل فعال بين المكون الانحداري الذاتي التاريخي وقدرات التعرف على الأنماط القوية للشبكات العصبية. هذا الجمع يسمح بالنقاط العلاقات الخطية وغير الخطية في بيانات معدل التضخم في سورية، مما

يوفر توقعات أكثر دقة وموثوقية، خاصة في ظل ظروف عدم اليقين. يقدم الإطار الرياضي الموصوف فهماً واضحاً لكيفية استخدام القيم السابقة وتحويلها من خلال الشبكة للتنبؤ بالقيم المستقبلية، مما يجعله أداة قوية لتحليل السلاسل الزمنية والتنبؤ بها.

2-9 نموذج بوابة الوحدة المتكررة Gated recurrent unit:

في حالة البيانات عالية التردد (يومية أو أسبوعية) تظهر مشكلة التسلسل إلى التسلسل Sequence to Sequence والتي من خلالها يصعب على النماذج السابقة NNAR بالتنبؤ بشكل صحيح بسبب عدم القدرة على الربط بين نقاط البيانات لفترات طويلة. في نماذج NNAR يقل التدرج $\frac{\partial L}{\partial \theta}$ (حيث L هو التلاشي و θ هي المعلمة)، بشكل أسي كلما انتشر إلى الخلف عبر الزمن. هذا يجعل من الصعب تعلم التبعيات طويلة الأمد. تعالج وحدات GRU هذه المشكلة باستخدام آليات التوجيه من خلال بوابة التحديث وإعادة الضبط التي تتحكم في تدفق المعلومات وتحافظ على التدرجات على التدرجات بشكل أكثر فعالية.

تعتبر وحدات GRU حاسمة في مشاكل التسلسل إلى التسلسل لأنها تستطيع توضيح التبعيات في تسلسلات حيث تكون الفجوة بين المعلومات ذات الصلة والنقطة المطلوبة كبيرة. وحدات البوابة المتكررة (GRU) هي نوع من الشبكات العصبية المتكررة (RNN) التي تكون فعالة بشكل خاص في معالجة تسلسل البيانات. يعمل نموذج GRU باستخدام بوابتين: بوابة إعادة الضبط وبوابة التحديث كما الشكل (2):



الشكل (2): معمارية نموذج GRU.

تحدد هذه البوابات مقدار المعلومات الماضية التي يجب تمريرها إلى المستقبل. تشمل المعادلات التي تحدد GRU (Alqahtani et al., 2023):

■ بوابة إعادة الضبط Reset Gate:

$$r_t = \sigma(X_t W_{xr} + H_{t-1} W_{hr} + b_r) \quad (6)$$

■ بوابة التحديث Update Gate:

$$z_t = \sigma(X_t W_{xz} + H_{t-1} W_{hz} + b_z) \quad (7)$$

■ حالة المرشح المخفية Candidate Hidden State:

$$\tilde{H}_t = \tanh(X_t W_{xh} + (r_t \odot H_{t-1}) W_{hh} + b_h) \quad (8)$$

■ الحالة المخفية النهائية Final Hidden State:

$$H_t = z_t \odot H_{t-1} + (1 - z_t) \odot \tilde{H}_t \quad (9)$$

في هذه المعادلات، يمثل (σ) الدالة السينية، ويشير (\odot) إلى الضرب حسب العناصر، و (\tanh) هي دالة ظل الزائدية. (W) و (b) هما الأوزان والتحييزات، على التوالي، وتتوافق العناصر السفلية (r) و (z) و (h) مع حالة إعادة التعيين والتحديث والحالة المخفية للمرشح.

يتكون نموذج Seq2Seq عادةً من: Encoder: يعالج سلسلة المدخلات ويلخص المعلومات في متجه السياق. Decoder: يستخدم متجه السياق لتوليد سلسلة المخرجات. بالنسبة لـ GRU، يتكون مدخلات ومخرجات التشفير كلاهما من وحدات GRU:

$$h_t^{enc} = \text{GRU}_{enc}(x_t, h_{t-1}^{enc}) \quad (10)$$

$$h_t^{dec} = \text{GRU}_{dec}(y_t, h_{t-1}^{dec}) \quad (11)$$

حيث c هو متجه السياق الذي يولده المشفر. لتدريب، نستخدم 90% من البيانات (حوالي 20700 مشاهدة) و 10% للاختبار (2300 مشاهدة). يتم تدريب النموذج لتقليل خطأ التنبؤ، الذي يُقاس عادةً باستخدام متوسط مربع الخطأ (MSE) لمهام الانحدار.

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (12)$$

حيث y_i العائد الفعلي و \hat{y}_i العائد المتوقع، نموذج GRU مناسب بشكل جيد لتنبؤ بيانات السلاسل الزمنية مثل عوائد سوق دمشق للأوراق المالية. باستخدام آليات التوجيه، يمكنه التعامل بفعالية مع التبعيات طويلة الأمد في البيانات. يعزز نهج Seq2Seq قدرته على تحويل سلاسل المدخلات إلى سلاسل المخرجات، مما يجعله خياراً قوياً للتنبؤ بالسلاسل الزمنية المالية. قدرة GRU على إدارة تدفق التدرج من خلال بوابات التحديث وإعادة الضبط تسمح لها بالنقاط التبعيات طويلة الأمد بشكل فعال. هذا يجعل وحدات GRU مناسبة بشكل خاص للبيانات عالية التردد في مشاكل Seq2Seq، حيث تواجه النماذج التقليدية مثل NNAR صعوبة بسبب مشكلة تلاشي التدرج. تضمن الصياغة الرياضية وآليات التوجيه في GRU أن يتم الاحتفاظ بالمعلومات ذات الصلة ونقلها عبر التسلسلات الطويلة، مما يثبت ميزتها في مثل هذه السياقات.

3-9 تقنية الغابة العشوائية Random Forest:

الغابة العشوائية هو أسلوب تعلم جماعي يستخدم لحل مشاكل الانحدار والتصنيف. حيث يعمل عن طريق إنشاء أشجار قرار متعددة أثناء مرحلة التدريب وإخراج الفصل الذي يمثل وضع الفئات (التصنيف) أو متوسط التنبؤ (الانحدار) للأشجار الفردية. يتمثل عمل خوارزمية الغابة العشوائية من خلال: محاكاة Bootstrap sampling يتم رسم عينات تمهيدية عشوائية من بيانات التدريب (Palimkar et al., 2022; He et al., 2022).

بالنظر إلى مجموعة بيانات $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ التي تحتوي على n من الملاحظات حيث يمثل x_i ميزات الإدخال و y_i المتغير المستهدف الثنائي (0 أو 1). يتم توليد عينة bootstrap من مجموعة البيانات الأصلية D حيث $b=1,2,3,\dots,B$. لكل عينة D_b Bootstrap نقوم بتدريب شجرة قرار T_b باستخدام مجموعة فرعية من الميزات، عند كل عقدة، يتم اختيار أفضل ميزة x_j من مجموعة فرعية عشوائية من الميزات لتقسيم البيانات، حيث $j \in \{1,2,\dots,m\}$ و m هو عدد الميزات في المجموعة الفرعية. يتم اختيار أفضل ميزة ونقطة تقسيم لتحقيق أقصى ربح معلومات (أو تقليل شوائب Gini أو الإنتروبيا). ربح المعلومات للانقسام s الذي يقسم العقدة t إلى العقتين الفرعيتين t_L و t_R هو:

$$\Delta G(s, t) = G(t) - \left(\frac{N_{tL}}{N_t} G(t_L) + \frac{N_{tR}}{N_t} G(t_R) \right) \quad (13)$$

بعد تدريب B شجرة، يتم عمل التنبؤات لملاحظة جديدة x عن طريق تجميع التنبؤات من جميع الأشجار. كل شجرة T_b تعطي تنبؤاً بالتصنيف $\hat{y}_b \in \{0,1\}$ يتم تحديده بواسطة التصويت بالأغلبية:

$$y = \text{mode}\{\hat{y}_1, \hat{y}_2, \dots, \hat{y}_B\} \quad (14)$$

يتم تحديد أهمية ميزة x_j من خلال إجمالي النقص في شوائب العقدة (مثل شوائب Gini) المنسوبة إلى الانقسامات باستخدام x_j عبر جميع الأشجار. لكل شجرة T_b يتم حساب النقص في الشوائب ΔG_j^b للميزة x_j :

$$\Delta G_j^b = \sum_{t \in \text{nodes using } x_j} \Delta G(s_t, t) \quad (15)$$

الأهمية الإجمالية للميزة x_j هي متوسط النقص في الشوائب عبر جميع الأشجار:

$$\text{Importance}(x_j) = \frac{1}{B} \sum_{b=1}^B \Delta G_j^b \quad (16)$$

ولتقييم التنبؤ في حالة التصنيف يتم استخدام منحني إشارات المستقبل ROC. يتم رسم المنحنى بوضع معدل الإيجابيات الصحيحة (TPR) على المحور الرأسي ومعدل الإيجابيات الخاطئة (FPR) على المحور الأفقي. المعادلة الرياضية لمنحنى ROC هي:

$$ROC = \frac{TPR}{FPR} \quad (17)$$

حيث $TPR = \frac{TP}{TP+FN}$ هو نسبة الحالات الإيجابية التي تم تصنيفها بشكل صحيح. وأن $FPR = \frac{FP}{FP+TN}$ هو نسبة الحالات السلبية التي تم تصنيفها بشكل خاطئ. لحساب AUC من منحني ROC، أحد الطرق الشائعة هو استخدام قاعدة (Trapezoidal rule)، والتي تقوم بتقسيم المنحنى إلى مجموعة من المستطيلات والمثلثات وحساب مساحة كل شكل وجمعها. الصيغة العامة لقاعدة تريبيز هي:

$$AUC = \sum_{i=1}^{n-1} \frac{f(x_i) + f(x_{i+1})}{2} \times (x_{i+1} - x_i) \quad (18)$$

حيث n هو عدد النقاط في منحني ROC، و x_i و $f(x_i)$ هي قيم FPR و TPR على التوالي في النقطة i . وبالنسبة للانحدار يتم استخدام متوسط مربع الخطأ الذي يعطى من خلال المعادلة:

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (19)$$

تهدف الغابات العشوائية إلى تحسين تباين الأشجار الفردية عن طريق حساب متوسط الأشجار المتعددة، مما يقلل من خطر الإفراط في التجهيز ويسمح للنموذج بتعميم البيانات غير الواضحة بشكل أفضل. تضمن العشوائية في اختيار الميزات لكل شجرة أن مجموعة الأشجار غير مترابطة، مما يجعل التنبؤ بالمتوسط أكثر قوة.

4-4- التنبؤ عالي الأبعاد ودالة الخسارة Lasso:

يتعمق القسم المقدم في تحديات وحلول التنبؤ الاقتصادي في بيئة بيانات عالية الأبعاد. حيث تحتوي مجموعات البيانات الحديثة (النصية والكمية) على أعداد كبيرة من المتنبئات المحتملة، مما يؤدي إلى مشاكل في التجهيز والتعدد الخطي Overfitting and Multicollinearity. هناك العديد من المحاولات التي تم فيها اقتراح معالجة هذه المشكلة على البيانات السورية في ظل الأبعاد العالية لكن مع ارتفاع عدد المتغيرات (أكثر من 10) تصبح هذه الأساليب غير دقيقة (Mansour et al., 2021; Alakkari, 2023).

على الرغم من أن مقدر المربعات الصغرى العادي يتمتع بالعديد من الصفات الجيدة، مثل عدم التحيز، إلا أنه يمكن أن يكون له تقلبات كبيرة في بعض المواقف. تقديرات المربعات الصغرى معرضة بدرجة كبيرة للأخطاء العشوائية وقد يكون لها تباين كبير، على سبيل المثال، إذا كانت البيانات تحتوي على العديد من المتغيرات المترابطة أو عدد من المتغيرات أكبر من طول مجموعة البيانات أي $N > T$ (غالباً ما يشار إليها بالمشكلة الكبيرة والصغيرة)، يمكن استخدام التنظيم لتقليل هذا التباين عن طريق إنشاء التحيز وخفض الخطأ الإجمالي. تعمل تقنيات الانحدار المُقَيّد *Penalized regression*، عن طريق تقليل أحجام متغيرات النموذج. إن إضافة مكون قيود إلى دالة التكلفة القياسية للانحدار الخطي هو الإجراء العادي:

$$J = \frac{1}{2m} \sum_{i=1}^m (y_i - \beta_0 - \sum_{j=1}^p x_{ij}\beta_j)^2 + \lambda \left[\frac{(1-\alpha)}{2} \sum_{j=1}^p \beta_j^2 + \alpha \sum_{j=1}^p |\beta_j| \right] \quad (20)$$

تتحول المعادلة إلى نموذج انحدار *Lasso* بناءً على قيمة الحد الجزائي. يتم تحديد تأثير القيود من خلال حجم معلمة القيود، $\lambda \geq 0$. في حالة تحديد قيمة "كبيرة" لـ λ ، يتم تقليل دالة التكلفة هذه إلى الحد الأدنى:

$$\min_{\beta} J \quad (21)$$

سيؤدي ذلك إلى انخفاض مستويات β أو ربما تصل إلى الصفر. وفقاً لذلك أقل تعقيداً وأقل عرضة للتركيب الزائد هو النموذج الذي يحتوي على معاملات أقل (أو صفر).

تعمل أساليب الانحدارات المنتظمة مثل *Lasso* و *Sparse-Group Lasso (sg-Lasso)* على تقليل معاملات الانحدار نحو الصفر، مما يقلل التباين ويمنع التجهيز الزائد. عامل الانكماش والتحديد المطلق الأقل *Least Absolute Shrinkage and Selection Operator (Lasso)* هو أسلوب انحدار يعمل على تقليل مجموع الأخطاء المربعة مع إضافة مصطلح قيود *L1* (Reid et al., 2016):

$$J = \frac{1}{2m} \sum_{i=1}^m (y_i - \beta_0 - \sum_{j=1}^p x_{ij}\beta_j)^2 + \lambda \sum_{j=1}^p |\beta_j| \quad (22)$$

لا يوجد حل تحليلي لأن حد القيود هو مجموع القيم المطلقة، مما يجعل الإجابة غير خطية. من الضروري حل معادلة *Lasso* كميًا. قد تقترب المعاملات من 0، من المهم ملاحظة أن *Lasso* عادةً ما يقلل من بعض المتغيرات الأخرى في المجموعة ويفضل المتغيرات المرتبطة بها (Ranstam and Cook, 2018). تشجع قيود *L1* على التناثر *Sparse*، مما يعني أن بعض المعاملات تصبح صفرًا تمامًا، مما يؤدي إلى إجراء اختيار متغير بشكل فعال. لحساب الحالات الخاصة لنموذج *Lasso*. نستكشف الشبكة عبر العديد من معاملات قيود *Lambda* ونقدر معلمة قيود *Lambda* واحدة في كل مرة. من أجل تحديد المعلمة ذات أقل معدل خطأ، فإننا نستخدم بالإضافة إلى ذلك طرق التحقق المتبادل *Cross Validation*. بعد التقدير، نعرض المزيد من إحصائيات التحقق والتشخيص، والرسوم البيانية لتطور المعامل فيما يتعلق بمعلمة القيود، وطرق العرض المخصصة لجداول المعامل والبيانات الموجزة الأخرى (Hastie et al., 2009; Friedman et al., 2010).

5-9 التحقق المتقاطع للسلاسل الزمنية *Time Series Cross Validation*:

يجب تحديد معلمة ضبط واحدة أو أكثر للتطبيق العملي لخوارزميات تعلم الآلة. من المعتاد استخدام التحقق المتقاطع *K-fold* للبيانات، والذي قد يكون أو لا يكون مناسباً لبيانات السلاسل الزمنية. وفقاً لـ (Bergmeir et al., 2018)، لا تزال تقنية التحقق المتقاطع التقليدية *K-fold* مناسبة لنماذج الانحدار الذاتي التي تحتوي على أخطاء.

ومع ذلك، يفشل التحقق المتبادل التقليدي مع الأخطاء المرتبطة بسبب الارتباط بين عينات التدريب والاختبار، كما هو الحال عندما يكون للانحدار تفسير إسقاطي فقط بسبب سوء التحديد. من أجل منع ذلك، ننظر أولاً إلى الارتباط التلقائي للأخطاء ونختار مجموعة من تقنيات التحقق المتبادل، ونستخدم:

Leave One Out: باستثناء الاحتفاظ بمجموعة اختبار بالحجم P واستخدام البيانات المتبقية كمجموعة تدريب، فإن هذا مشابه لـ K -Fold. يتم تكرار هذا الإجراء لكل من المجموعات المتبقية ويتم استخدام الملاحظات المحذوفة لتقييم النموذج (Bergmeir et al., 2018):

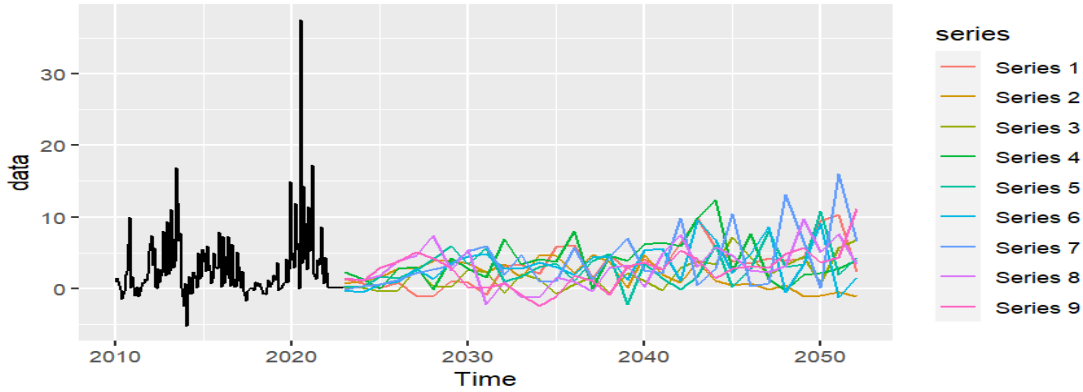
$$CV(\lambda) = \frac{1}{T} \sum_{t=1}^T \ell(y_t - \hat{f}_{\lambda,-t,l}(x_t)) \quad (23)$$

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (24)$$

حيث ℓ هي دالة الخسارة MSE_i و $\hat{f}_{\lambda,-t,l}(x_t)$ تمثل قاعدة التعلم الآلي المستخدمة في التنبؤ.

10- النتائج:

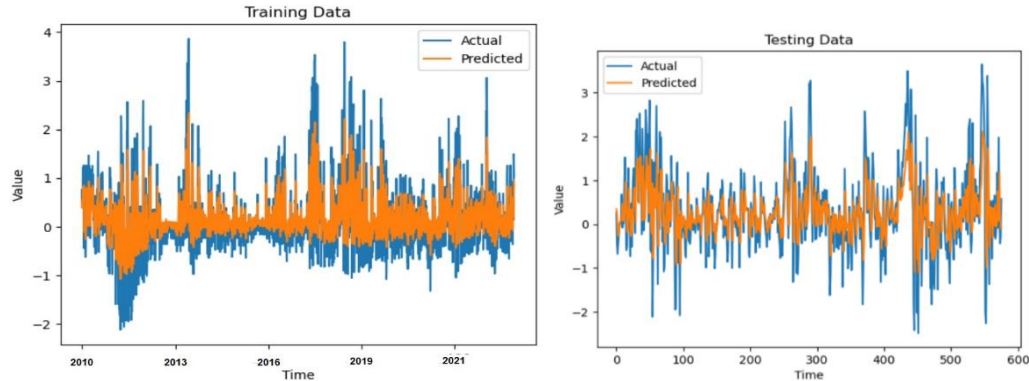
من أهم المشاكل في إطار بيانات الاقتصاد السوري هو التأخر في جمع البيانات ونشرها بسبب الإمكانيات المادية المحدودة، يمكن من خلال هذه البيانات الحصول على رؤى طويلة المدى للتنبؤات مع معالجة مشكلة التدرج المتلاشي، مع بيانات معدل التضخم الشهرية والمتوفرة حتى نهاية 2020 يمكن تقديم رؤى لتطور المتغير حتى عام 2050 من خلال الشكل (3) نحصل على نتائج التنبؤات لنموذج NNAR:

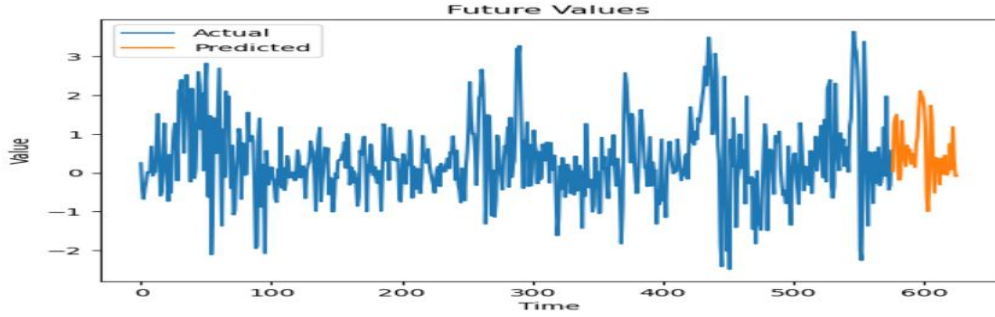


الشكل (3): التنبؤ بمعدل التضخم في سورية حتى عام 2050 باستخدام محاكاة الشبكات العصبية (تشير مجموعة خطوط التنبؤ إلى سيناريوهات التنبؤ المولدة من محاكاة الشبكة العصبية والتي تشكل فترة عدم التأكد) - (استقرار الفترة الملاصقة لفترة التنبؤ هو أن النموذج يتضمن تقنيات لتقليل التباين مقابل بعض التحيز في فترة التدريب).

ولحل مشكلة التسلسل إلى التسلسل من خلال نموذج GRU يتم التنبؤ بعوائد أسعار سوق دمشق

للأوراق المالية من خلال الشكل (4):



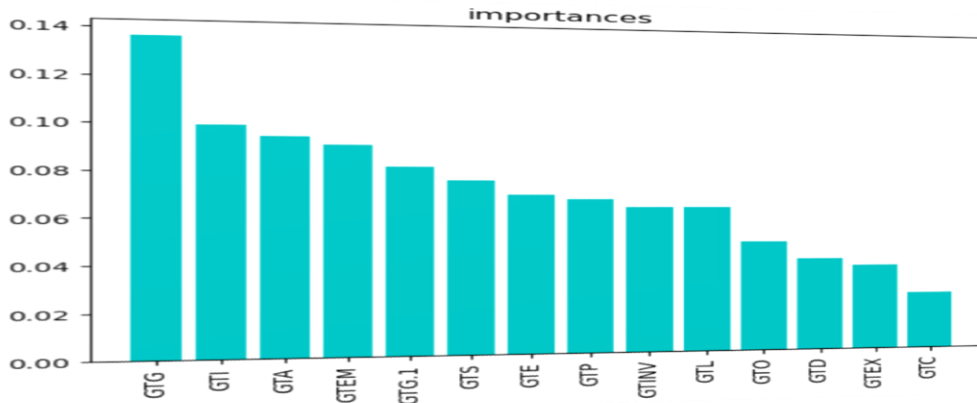


الشكل (4): نتائج تدريب النموذج واختباره والتوقع على بيانات مؤشر سوق دمشق للأوراق المالية (أيام).

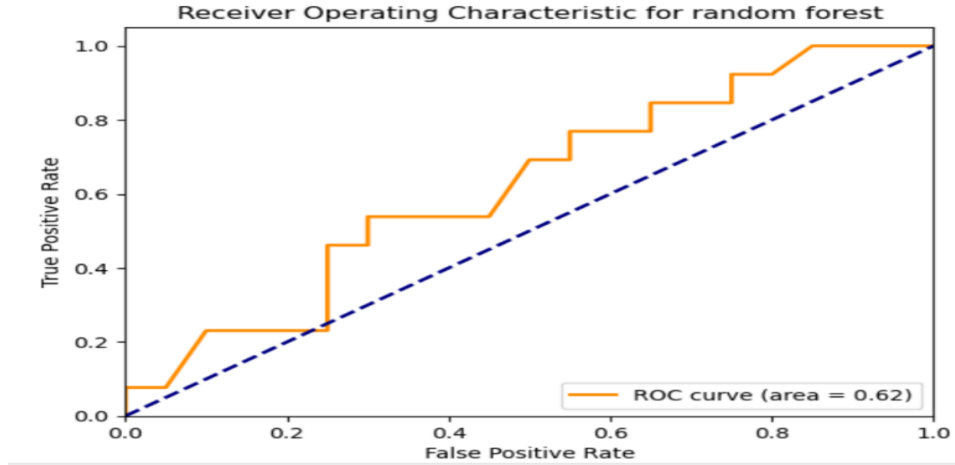
لحل مشاكل التصنيف واتخاذ القرار الأمثل بالاعتماد على عدد كبير من المتغيرات نعمل على تطبيق تقنية الغابة العشوائية على بيانات الاقتصاد السوري، نعمل على جمع مجموعة من البيانات النصية المتوفرة من خلال بحث Google Trend والتي تتضمن أنماط البحث عن كلمات مفتاحية تتعلق بالاقتصاد السوري بالنسبة لجميع الأفراد المقيمين في سورية. يوفر Google Trend فهرساً لحجم البحث (SVI_{st}) (Search Volume Index) في Google منذ عام 2004، يقيس حجم البحث عن مصطلح ما (أو مجموعة مصطلحات) S في منطقة ما خلال الفترة t (Chu and Marron, 1991):

$$SVI_{st} = \frac{sv_{st}}{sv_{Gt} \times MSV_{[0,T]}} \times 100 = \frac{sv_{st}}{sv_{Gt} \times \max_{t=[0,T]} [sv_{st}/sv_{Gt}]} \times 100 \quad (18)$$

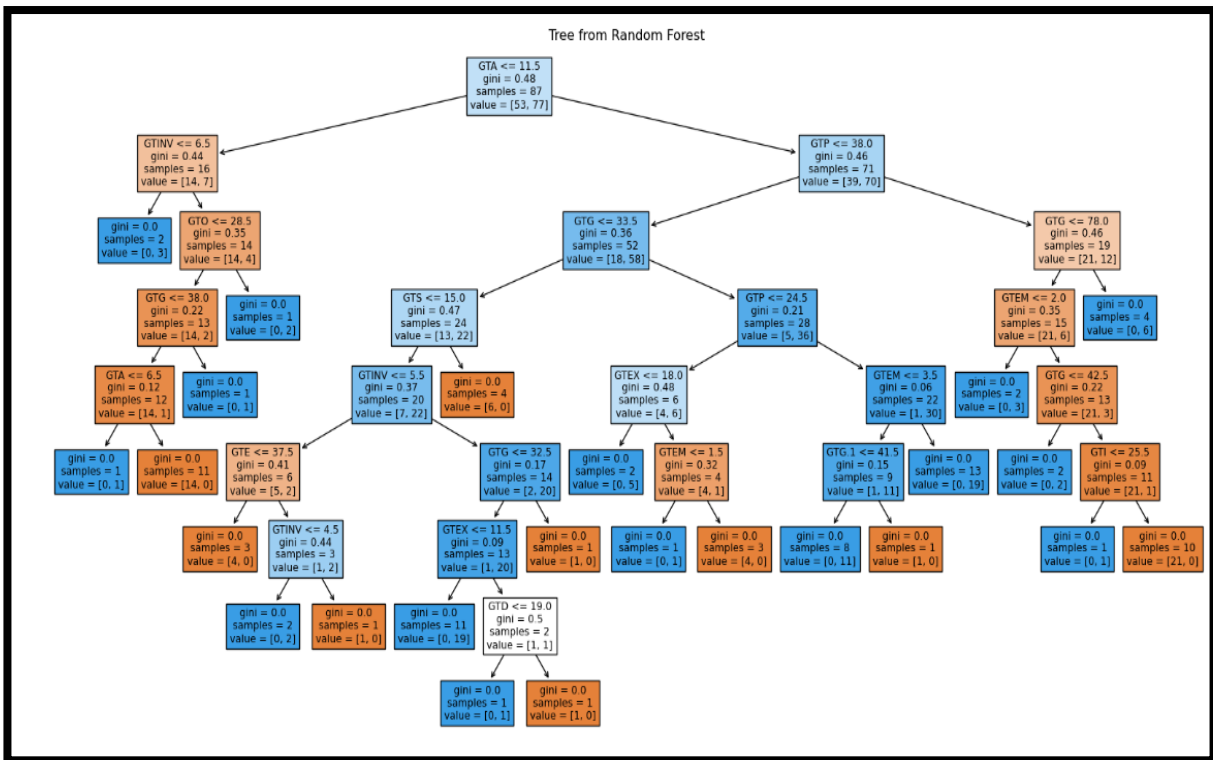
حيث تُمثل sv_{st} عدد عمليات البحث عن المصطلح S خلال فترة t (تردد المشاهدات شهري)، يمثل sv_{Gt} عدد عمليات البحث في Google خلال نفس الفترة، وبالتالي SVI_{st} تُمثل سلسلة زمنية تأخذ قيماً بين 0 و 100 بحيث يتم قياسها بواسطة الحد الأقصى لقيمة sv_{st}/sv_{Gt} خلال الفترة الزمنية المحددة $[0, T]$. نستخدم الكلمات المفتاحية التي ترتبط بالاقتصاد السوري (الذهب، سعر الصرف، الصناعة، الزراعة، الاستثمار، الاستهلاك، السوق، الليرة السورية، التجارة، التوظيف، الأسهم) ونحصل على النتائج من خلال الشكل (5-6-7).



الشكل (5): الأهمية النسبية للمتغيرات في قدرتها على التصنيف باستخدام الغابة العشوائية.



الشكل (6): منحنى إشارات المستقبل ROC لتقييم جودة النموذج.



الشكل (7): نتائج خوارزمية الغابة العشوائية.

في التصميم الأخير لتحقيق التكامل مع التقنيات السابقة نعمل على استخدام 25 متغير من الاقتصاد السوري للتنبؤ بأسعار المستهلك من خلال خوارزمية Lasso كما هو موضح في الجدول (1). بالتطبيق باستخدام انحدار Lasso مع فرض قيود L_1 نحصل على النتائج في الجدول (2) حيث يعرض هذا الجدول نتائج تطبيق انحدار Lasso على البيانات الاقتصادية السورية بهدف التنبؤ بمؤشر أسعار المستهلك (CPI) كمتغير تابع.

الجدول (2): نتائج انحدار Lasso.

Dependent Variable: CPI	
Method: Elastic Net Regularization	
Sample (adjusted): 2010M01 2022M12	
Included observations: 156 after adjustments	

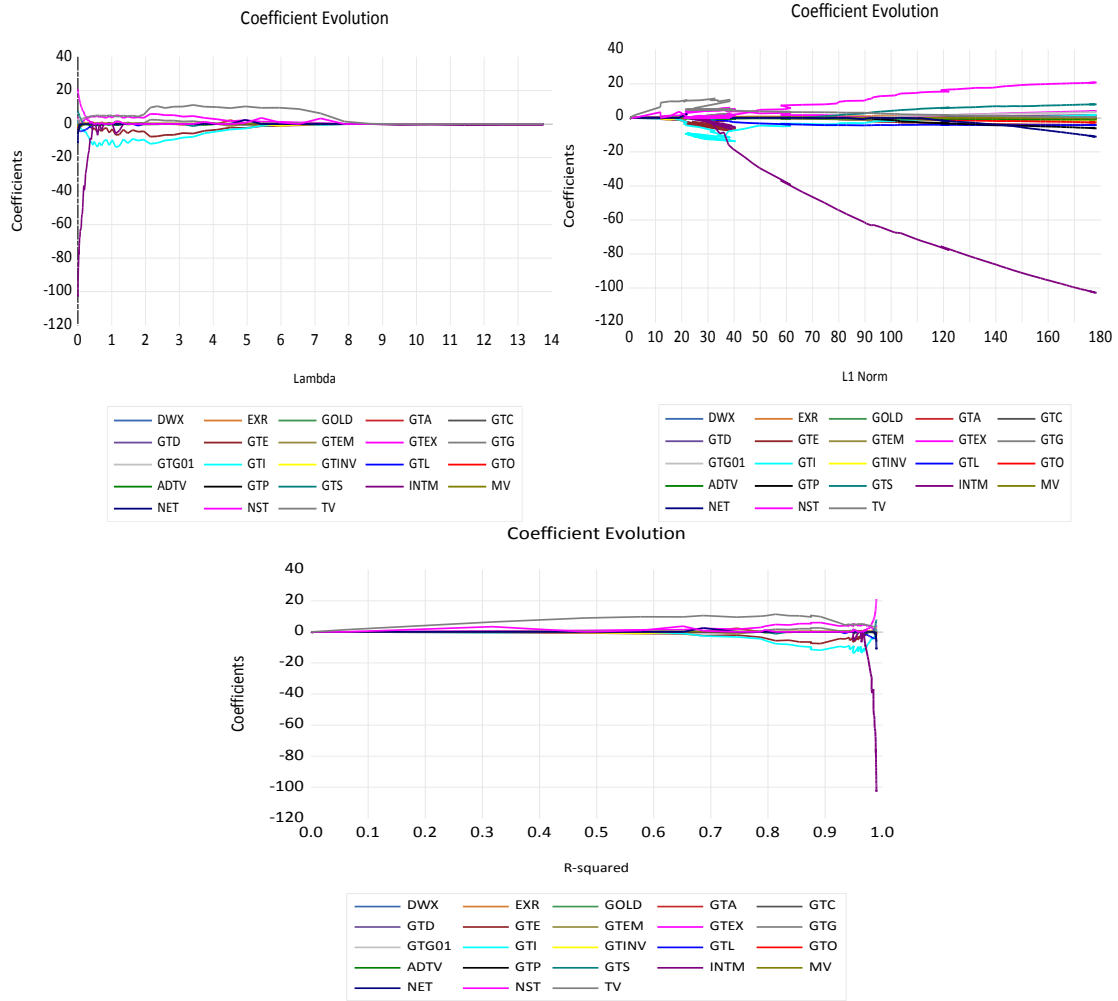
Penalty type: Lasso (alpha = 1)			
Lambda at minimum error: 0.00219			
Regressor transformation: L1			
Cross-validation method: K-Fold (number of folds = 5), rng=kn,			
seed=113837611			
Selection measure: Mean Squared Error			
	(minimum)	(+ 1 SE)	(+ 2 SE)
Lambda	0.00219	0.3653	1.017
Variable	Coefficients		
DWX	0.056689	0.046920	0.047876
EXR	1.205668	0.563508	0.491043
GOLD	0.005332	0.005038	0.004611
GTA	-1.135077	1.705298	-1.00E-07
GTC	-0.717428	-1.69E-05	0.000000
GTD	1.537490	-9.38E-05	-1.63E-08
GTE	-4.082311	-2.158005	-2.973473
GTEM	-1.517255	-8.48E-05	3.44E-08
GTEX	3.865007	4.567998	3.804785
GTG	0.150269	1.65E-05	5.30E-08
GTG01	3.275186	-7.07E-06	0.000000
GTI	1.421722	-11.24895	-9.482813
GTINV	-1.159544	1.85E-05	-1.40E-07
GTL	-4.300252	-1.555131	-0.001688
GTO	-2.600120	-8.52E-05	0.000000
ADTV	-0.854073	-0.555378	-0.163117
GTP	-5.849779	-1.83E-05	0.000000
GTS	7.841092	0.000000	0.000000
INTM	-101.1000	-8.474252	8.10E-07
MV	-0.201825	0.257493	0.266377
NET	-10.00805	-0.000487	2.27E-06
NST	20.57280	0.277607	2.22E-07
TV	0.519753	4.719580	4.971590
C	1207.958	641.6661	564.5103
d.f.	23	22	18
L1 Norm	1381.935	677.8020	586.7177
R-squared	0.989914	0.970510	0.958709

المصدر: اعداد الباحثان بالاعتماد على بيئة بايثون - جوبيتر.

(Method): Elastic Net Regularization تظهر المعلومات الواردة في نموذج التقدير ما يلي: الطريقة

with Lasso (alpha = 1) والعينة (Sample): البيانات المستخدمة في التحليل من كانون الثاني 2010 إلى كانون الأول 2022 (شهرية). والمشاهدات (Observations): 156 مشاهدة بعد التعديلات. ونوع القيود (Penalty type): Lasso (L1 norm) و (Lambda at minimum error): 0.00219 (قيمة Lambda التي تحقق أقل خطأ في التنبؤ). وتحويل الانحدار (Regressor transformation): L1 (مجموع القيم المطلقة للمعاملات). والتحقق المتبادل K-Fold (Cross-validation method) مع 5 طيات (لتقييم أداء النموذج). وقياس الاختيار (Selection measure): متوسط مربع الخطأ (MSE) لتقييم دقة النموذج.

يعرض الشكل (8) تطور معاملات النموذج وتقدم الرسوم البيانية المضمنة فيه معلومات قيمة حول سلوك معاملات انحدار Lasso وأداء النموذج في التنبؤ بمؤشر أسعار المستهلك (CPI) في سورية. حيث: الشكل أعلى اللوحة على اليمين: تطور المعاملات بالنسبة للمقياس L1.

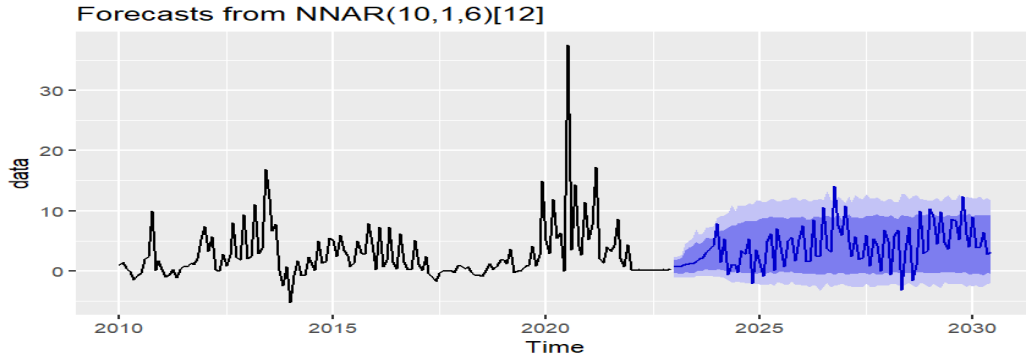


الشكل (8): تطور معاملات نموذج انحدار Lasso.

الشكل أعلى اللوحة على اليسار: تطور المعاملات بالنسبة لـ λ : يعرض هذا الشكل البياني كيف تتغير قيم المعاملات بتغير قيمة معامل الضبط λ . الشكل أسفل اللوحة: تطور المعاملات بالنسبة لـ R^2 : يعرض هذا الشكل البياني كيف تتغير قيم المعاملات بتغير قيمة معامل التحديد R^2 .

11- مناقشة النتائج:

يبين الشكل (9) نتائج التنبؤات مع محاكاة الشبكات العصبية لسنايروهات التنبؤ والتي تمثل محاكاة لفترات عدم التأكد لهذه التنبؤات وفق الشكل (9):

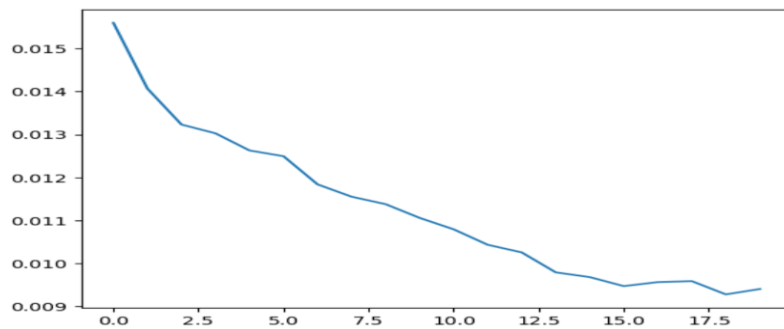


الشكل (9): التنبؤات بمعدل التضخم الشهري في سورية مع تضمين حالات عدم التأكد.

نلاحظ بأنه تم تضمين فترات عدم التأكد مع التنبؤات المركزية لمعدل التضخم وبالتالي مع تنبؤات بمعدل 3.14% للتضخم في نهاية 2030 مع هامش عدم تأكد (-0.49%, 9.22%) مع اتباع نفس السياسات الحالية نلاحظ من خلال الجدول (2) مع مقارنة بيانات الاختبار مع البيانات الفعلية أن التنبؤات الفعلية خلال عام 2022 تقع ضمن فترة عدم التأكد للتنبؤات باستخدام النموذج، وبمقارنة بيانات الاختبار مع البيانات الفعلية نقوم بحساب نسبة متوسط الخطأ المطلق وفق المعادلة:

$$MAPE = \sqrt{\frac{\sum_{t=1}^n (\hat{y}_t - y_t)^2}{n}} = 3.6276 \quad (25)$$

نلاحظ أن نسبة الخطأ في التنبؤ 3.7% وبالتالي تبلغ نسبة الدقة فيها 96.3% وهو ما يشير إلى أنه يمكن الوثوق بالنتائج في ظل الظروف الحالية. ولكن في حالة البيانات عالية التردد (يومية أو أسبوعية) تظهر مشكلة التسلسل إلى التسلسل Sequence to Sequence والتي من خلالها يصعب على النماذج السابقة بالتنبؤ بشكل صحيح بسبب عدم القدرة على الربط بين نقاط البيانات لفترات طويلة وهو ما ينطبق على بيانات عوائد سوق دمشق للأوراق المالية اليومية، بالدمج مع دالة Rule نقوم بتدريب النموذج على بيانات عوائد سوق دمشق للأوراق المالية اليومية خلال الفترة 2010-2024 ويتم الحصول على 90% من البيانات للتدريب (2300 مشاهدة) حيث بلغ عدد المرات التي ستعمل فيها خوارزمية التعلم من خلال مجموعة بيانات التدريب بأكملها epochs من خلال الشكل (10) نجد أن النموذج يُظهر تحسناً مع زيادة عدد (epochs)، حيث ينخفض معدل الخسارة (loss) مع تقدم الفترات:



الشكل (10): فعالية تدريب النموذج على البيانات epochs.

هذا يشير إلى أن النموذج يتعلم ويحسن أداءه بمرور الوقت. وبالتالي نحصل على نتائج التدريب والاختبار والتنبؤ بالفترات المستقبلية ل 50 يوم اعتباراً من 7/3/2024 للنموذج على البيانات وفق الشكل (4). حيث أن معامل

الجذر التربيعي لمربع الخطأ في التنبؤات بلغ 0.85 مما يدل على تقارب القيم الفعلية والمقدرة. يمكن هذا الموضوع المستثمرين والمتداولين على الحصول على معلومات مستتيرة حول اتجاهات تطور السوق مما يساعدهم في اتخاذ قرار الاستثمار في الوقت المناسب.

تم تضمين 14 متغيراً تدل على أنماط البحث على الانترنت حول الاقتصاد السوري في خوارزمية الغابة العشوائية ونقوم باستخدامها للتنبؤ بالمتغير الهدف المعروف بالتصنيف الآتي: 1: اتخاذ القرار الاستثماري، والتي تعطى قيمة مقابل الحصول على أرباح في سوق الأوراق المالية و0: عدم اتخاذ القرار الاستثماري تعطى لكل حالة مقابلة للخسائر في السوق المالية. يتم تطبيق خوارزمية الغابة العشوائية ونحصل على النتائج الموضحة في الشكل (7). حيث نلاحظ أن أنماط البحث عن أسعار الذهب (GTG) هي الأكثر أهمية للتنبؤ باتخاذ القرار الاستثماري في السوق المالي وبالنسبة لجميع المتغيرات فإنها تعطي نسبة قرار صحيحة لاتخاذ القرار الاستثماري في السوق المالي بنسبة 62% وفق الشكل (6) حيث نلاحظ من الشكل أن جميع التنبؤات تقع في القسم الإيجابي والمقبول لمنحنى إشارات المستقبل ROC وبالتالي مع كل تحديث للمعلومات هذه يمكن معرفة اتجاه السوق سواء كان صعوداً أو هبوطاً وبالتالي اتخاذ القرار الاستثماري الأكثر ملائمة في الوقت المناسب حيث يظهر من الشكل (7) شجرة قرار (نتيجة) نموذج الغابة العشوائية تُستخدم لتحليل البيانات المتعلقة بالاقتصاد السوري.

تستخدم الشجرة متغيرات مختلفة مع عتبات محددة لاتخاذ قرارات ثنائية، مما يؤدي إلى التنبؤ إما ب 1 (اتخاذ قرار استثماري من المتوقع أن يؤدي إلى ربح) أو 0 (عدم اتخاذ قرار استثماري من المتوقع أن يؤدي إلى خسارة). تحتوي الشجرة على عقد وفروع تشير إلى مسارات مختلفة بناءً على المعايير المحددة في كل عقدة. تحتوي كل عقدة على معلومات حول أحد المتغيرات الـ 14 المتعلقة بالاقتصاد السوري، مع عتبات محددة لاتخاذ القرارات. العقد ملونة: الأزرق للعقد التي تؤدي إلى التنبؤ ب 0 والبرتقالي للعقد التي تؤدي إلى التنبؤ ب 1. تعرض كل عقدة المتغير والعتبة التي يتم اختبارها، درجة نقاء جيني، عدد العينات المتأثرة، وتوزيع الفئة المتوقع. حيث، يمكننا رؤية عقدة تختبر متغير "GTG" وهو الأكثر أهمية نسبية بين المتغيرات مع عتبة أقل من أو تساوي 33.5. إذا كانت القيمة أقل من هذه العتبة، فإن الشجرة تتبع فرعاً يؤدي إلى التنبؤ ب 0، وهو ما يشير إلى عدم اتخاذ قرار استثماري بسبب التوقع بالخسارة. العقدة تُظهر أيضاً درجة نقاء جيني حيث أن:

$$G = 1 - \sum_{i=1}^n p_i^2 \quad (26)$$

حيث (G) هو مؤشر جيني، و (pi) هو احتمال وجود عينة من الفئة (i) في مجموعة البيانات المعينة، و (n) هو عدد الفئات. الفكرة هي أنه كلما كانت العينات ضمن عقدة معينة تنتمي إلى فئة واحدة، كان مؤشر جيني أقل، مما يعني أن العقدة نقية والتي تساوي من خلال الشكل 0.36 أي أن نقاء التنبؤ يبلغ درجة 64% وعدد العينات المتأثرة 52 وتوزيع الفئة المتوقعة لهذا القرار يتراوح بين 18 و58.

يعرض الجدول (2) لنتائج انحدار Lasso ثلاث مجموعات من المعاملات عند قيم Lambda مختلفة: (+2 SE), (+1 SE), (minimum). يمثلان قيم Lambda التي تكون أكبر بمرتين ومرتين من الانحراف المعياري للقيمة المثلى. وزيادة قيمة Lambda تؤدي إلى تقليص قيم المعاملات نحو الصفر، مما يزيد من بساطة النموذج ويقلل من التباين. عند قيمة Lambda الأكبر (+2 SE)، نلاحظ أن العديد من المعاملات أصبحت صفراً، مما يعني أن هذه المتغيرات تم استبعادها من النموذج. تشير قيمة المعامل إلى تأثير المتغير

على مؤشر أسعار المستهلك (CPI) مع ثبات باقي المتغيرات. المعاملات الموجبة تدل على علاقة طردية مع CPI (زيادة المتغير تؤدي إلى زيادة CPI) والمعاملات السالبة تدل على علاقة عكسية مع CPI (زيادة المتغير تؤدي إلى انخفاض CPI) المعاملات التي تساوي صفراً تعني أن المتغير ليس له تأثير كبير على CPI وتم استبعاده من النموذج.

EXR (سعر الصرف): معامل موجب وقيمه كبيرة نسبياً، مما يشير إلى أن ارتفاع سعر الصرف يؤدي إلى زيادة ملحوظة في مؤشر أسعار المستهلك.

GTS (اتجاه البحث عن الأسهم): معامل موجب وقيمه كبيرة، مما يعني أن زيادة البحث عن الأسهم على جوجل مرتبطة بارتفاع مؤشر أسعار المستهلك.

GTE (اتجاه البحث عن التعليم): معامل سالب، مما يشير إلى أن زيادة البحث عن التعليم على جوجل مرتبطة بانخفاض مؤشر أسعار المستهلك.

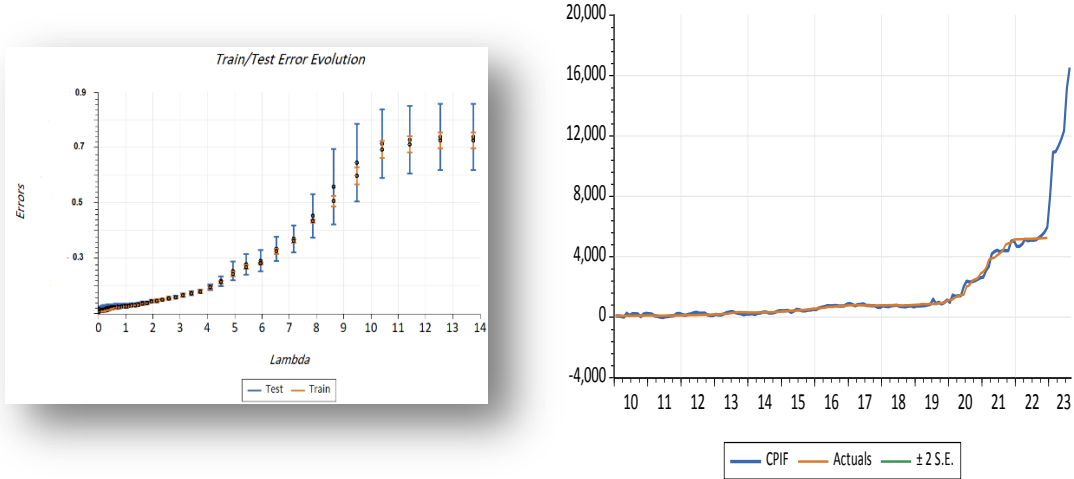
L1 Norm: مجموع القيم المطلقة للمعاملات، وهو مقياس للتناثر في النموذج.

R-squared: معامل التحديد الذي يقيس نسبة التباين في CPI التي يفسرها النموذج. القيمة 0.989914

عند Lambda الأمثل تشير إلى أن النموذج يفسر نسبة كبيرة من التباين في CPI.

يوضح الشكل البياني (8) كيف تتغير قيم معاملات الانحدار بتغير قيمة المقياس L1 (مجموع القيم المطلقة للمعاملات). مع زيادة قيمة L1، يتم تقليص المعاملات نحو الصفر، مما يزيد من بساطة النموذج ويقلل من التباين. يمكن ملاحظة أن بعض المعاملات تصبح صفراً عند قيم L1 معينة، مما يعني أن هذه المتغيرات تم استبعادها من النموذج. يسمح هذا الشكل البياني بتحديد مدى حساسية المعاملات للتغيرات في قيمة L1 ثم اختيار المناسب منها في النموذج. زيادة قيمة Lambda تؤدي إلى تقليص المعاملات نحو الصفر بشكل مشابه للشكل الأول. يمكن ملاحظة أن بعض المعاملات تصبح صفراً عند قيم Lambda معينة، مما يعني استبعاد هذه المتغيرات. يسمح هذا الشكل البياني بتحديد قيمة Lambda المثلى التي تحقق أفضل توازن بين الدقة والبساطة والتي تم اعتمادها في تقدير النموذج. مع زيادة قيمة R-squared، يتم تقليص المعاملات نحو الصفر، مما يشير إلى أن النماذج الأكثر دقة تميل إلى أن تكون أكثر بساطة.

يسمح هذا الشكل البياني بتحديد العلاقة بين دقة النموذج وتعقيده وفق أفضل درجة والتي تحقق أفضل تقارب بين القيم الفعلية والمتوقعة وما يبينه الشكل (11)، حيث يبين الشكل في الجهة اليمين تطور خطأ التدريب والاختبار بالنسبة ل Lambda: يوضح هذا الشكل البياني كيف يتغير خطأ التنبؤ في مجموعتي التدريب والاختبار بتغير قيمة Lambda:



الشكل (11): أخطاء النموذج التدريب والاختبار وتطور التنبؤ بأسعار المستهلك.

مع زيادة قيمة λ ، يزداد خطأ التدريب والاختبار، مما يشير إلى أن النماذج الأكثر بساطة تميل إلى أن تكون أقل دقة. يسمح هذا الشكل البياني بتحديد قيمة λ المثلى التي تحقق أفضل توازن بين خطأ التدريب والاختبار. وعدم وجود فجوة كبيرة بين خطأي التدريب والاختبار يشير إلى عدم وجود مشكلة التجهيز الزائد. الشكل في الجهة اليسرى: القيم الفعلية والمتوقعة لـ CPI يعرض هذا الشكل البياني القيم الفعلية والمتوقعة لـ CPI يمكن ملاحظة أن النموذج يتتبع الاتجاه العام لـ CPI بشكل جيد، ويتوقع أن يصل مؤشر أسعار المستهلك إلى حوالي 16500 نقطة مع شهر 9 من عام 2023.

يوفر انحدار Lasso أداة فعالة لاختيار المتغيرات المهمة وتقليل التباين في التنبؤ بمؤشر أسعار المستهلك في سورية. تؤثر العديد من العوامل الاقتصادية والاجتماعية على CPI، بما في ذلك سعر الصرف، اتجاهات البحث على جوجل، أسعار الذهب، وغيرها حيث يمكن استخدام النتائج لتطوير سياسات اقتصادية فعالة تستهدف العوامل الرئيسية المؤثرة على التضخم.

تُظهر نتائج هذه الدراسة إمكانية دمج تقنيات تعلم الآلة في تحليل البيانات الاقتصادية السورية المعقدة، مما يفتح آفاقاً جديدة لدعم القرارات الاقتصادية. وقد أظهرت الشبكات العصبية قدرةً فائقةً على نمذجة التطورات غير الخطية المعقدة بين المتغيرات الاقتصادية، كما يتضح من خلال دقة التنبؤ بمعدل التضخم على المدى الطويل، الذي فاق توقعات النماذج التقليدية. تُشير هذه النتائج إلى إمكانات تعلم الآلة في التغلب على قيود النماذج الاقتصادية الكلاسيكية التي تعاني من صعوبة التعامل مع البيانات غير الخطية، كما تلاحظ الدراسات (Bulut, 2018; Al-Akkari and Ali, 2022).

على الرغم من قلة البيانات المتاحة، فقد تمكن نموذج GRU من التنبؤ باتجاه تطور مؤشر سوق دمشق للأوراق المالية بشكلٍ دقيق، مما يوفر معلومات قيمة للمستثمرين. تُعد هذه النتيجة مُشابهة لما تمّ التوصل إليه في دراسات سابقة مثل (Ali et al., 2024) التي استخدمت تقنيات تعلم الآلة لتنبؤات أسعار الأسهم في أسواق أخرى، مما يُشير إلى فعالية هذه التقنيات في مختلف الظروف. وتُسلط نتائج الغابة العشوائية الضوء على الدور الكبير للبيانات النصية، المُستخلصة من Google Trends، في فهم اتجاهات السوق واتخاذ القرارات الاستثمارية، مما يؤكد على أهمية دمج مصادر البيانات غير التقليدية في التحليل الاقتصادي، كما تقترح الدراسات (Song and Choi, 2023; Borup and Schütte, 2022). وتُبرز نتائج انحدار Lasso

فعالية التقنية في التعامل مع التنبؤات عالي الأبعاد، مما يُمكن من تحديد المتغيرات الأكثر تأثيراً على مؤشر أسعار المستهلك، وتقليل التباين في التنبؤات. تُشابه هذه النتائج ما توصلت إليه الدراسات (Chandra and Manuaba, 2023; Babii, 2022; Liu et al., 2023) التي أظهرت قدرة انحدار Lasso على تحسين دقة التنبؤات الاقتصادية في بيانات البيانات عالية الأبعاد.

تفتح هذه النتائج آفاقاً جديدة لدراسات مستقبلية، حيث يمكن تطوير نماذج تنبؤية أكثر دقة للظواهر الاقتصادية الرئيسية مثل النمو الاقتصادي وسعر الصرف، مما يُمكن من اتخاذ قرارات سياسية أفضل وأكثر استنارة.

11- الاستنتاجات والتوصيات:

11-1 الاستنتاجات:

تُظهر الدراسة الحالية كيف يمكن لتقنيات التعلم الآلي أن تُستخدم كأدوات فعالة لتحليل البيانات الاقتصادية السورية المعقدة وتوفير رؤية قيمة لدعم القرارات الاقتصادية المرتبطة بضبط معدل التضخم بما يتماشى مع الظروف الحالية، ومعرفة العوامل الأكثر تأثيراً عند اتخاذ القرار الاستثماري أو عدم اتخاذه وبالتالي إمكانية اتخاذ قرارات لتخفيف حالة المخاطر في السوق، مع معرفة العوامل الأكثر تأثيراً في أسعار المستهلك وبالتالي ضبطه من خلالها أو تفعيل عوامل ذات أهمية غير مؤثرة. من خلال استخدام مجموعة متنوعة من الأساليب، بما في ذلك الشبكات العصبية، والغابات العشوائية، وانحدار Lasso، تمكنا من بناء نماذج تنبؤية دقيقة لمتغيرات اقتصادية مهمة مثل التضخم ومؤشر سوق دمشق واستخلاص أنماط وتنبؤات مهمة من البيانات الاقتصادية المتاحة.

أثبتت الشبكات العصبية فعاليتها في نمذجة التطورات غير الخطية المعقدة بين المتغيرات الاقتصادية، كما ينصح من خلال توقع معدل التضخم على المدى الطويل. كما تم استخدام نموذج GRU للتنبؤ باتجاه تطور مؤشر سوق دمشق للأوراق المالية حيث توفر هذه القدرة على التنبؤ معلومات قيمة لصانعي السياسات والشركات والأفراد في التخطيط الاستراتيجي واتخاذ القرارات الاقتصادية. علاوة على ذلك، أظهرت الغابات العشوائية قيمتها في تحليل البيانات النصية المستخرجة من Google Trends، مما يسلط الضوء على الدور المهم الذي يلعبه البحث عبر الإنترنت في فهم اتجاهات السوق واتخاذ القرارات الاستثمارية.

من ناحية أخرى، قدم انحدار Lasso حلاً قوياً للتحديات التي يفرضها التنبؤ عالي الأبعاد، مما يمكن من تحديد المتغيرات الأكثر تأثيراً على مؤشر أسعار المستهلك وتقليل التباين في التنبؤات. تسلط هذه النتائج الضوء على إمكانات التعلم الآلي في تحسين دقة التنبؤ الاقتصادي في بيئة بيانات عالية الأبعاد ومعقدة مثل سورية.

11-2 التوصيات:

بالنظر إلى النتائج، تقترح الدراسة زيادة دمج تقنيات التعلم الآلي في عمليات التخطيط والتحليل الاقتصادي في سورية. يمكن استخدام هذه التقنيات لتطوير نماذج تنبؤية أكثر دقة للظواهر الاقتصادية الرئيسية، مثل النمو الاقتصادي، والتضخم، وسعر الصرف. حيث تسهم هذه النماذج في اتخاذ قرارات سياسية أكثر استنارة وفعالية، مما يساعد على تعزيز المعلومات وتخفيف حالات عدم اليقين الاقتصادي مما يؤكد ضرورة إنشاء مركز متخصص للذكاء الاصطناعي حيث يهدف المركز إلى تطوير قدرات المكتب المركزي للإحصاء في سورية في تحليل البيانات باستخدام تقنيات التعلم الآلي، وتطوير نماذج تنبؤية للعديد من المتغيرات الاقتصادية.

12- المراجع:

1. مرهج، منذر. 2013. التنبؤ بقيم مؤشر سوق دمشق للأوراق المالية باستخدام الشبكات العصبية الاصطناعية واتجاهاتها. مجلة جامعة تشرين للبحوث والدراسات العلمية، المجلد 35، العدد 5.
2. الحميد، أيهم، خلف، أسهان، نقار، عثمان. 2020. تحسين القدرة على التنبؤ باتجاه حركة مؤشر سوق دمشق للأوراق المالية باستخدام أسلوب مقترح لبناء نموذج هجين بين الشبكات العصبونية الاصطناعية ARCH-ARIMA. مجلة جامعة حماه، المجلد 3، العدد 4.
3. يوسف، مظهر. 2024. التنبؤ بالأرقام القياسية لأسعار المستهلك استناداً لسعر الصرف في سورية. مجلة جامعة دمشق للعلوم الاقتصادية والسياسية، المجلد 40، العدد 3.
- المكتب المركزي للإحصاء (٢٠٢٢). المجموعة الإحصائية السورية ٢٠٢١. سورية: دمشق.

<http://www.cbssyr.sy>

- 1- Alakkari K. *Nowcasting GDP in Parsimony Data Environment Using Bayesian Mixed Frequency VAR (Empirical Evidence from Syria)*. Syrian Journal for Science and Innovation. 2023;1(2-Volume 1).
- 2- Al-Akkari Khder, Ali B. *Using Bayesian Inference for Stochastic Volatility Model (Empirical Evidence from Damascus Securities Exchange)*. Tishreen University Journal- Economic and Legal Sciences Series. 2022;44(3):11-32.
- 3- Ali B, Alakkari K, Abotaleb M, Mijwil MM, Dhoska K. *MLP and RBF Algorithms in Finance: Predicting and Classifying Stock Prices amidst Economic Policy Uncertainty*. Mesopotamian Journal of Big Data. 2024 May 11; 2024:48-67.
- 4- Alqahtani F, Abotaleb M, Subhi A, El-Kenawy S, Abdelhamid A, Alakkari K, et al. *A hybrid deep learning model for rainfall in the wetlands of southern Iraq*. Modeling Earth Systems and Environment. 2023;9(4):4295-4312.
- 5- Babii A. *High-dimensional mixed-frequency IV regression*. Journal of Business & Economic Statistics. 2022 Oct 2;40(4):1470-83.
- 6- Bergmeir C, Hyndman R, Koo B. *A note on the validity of cross-validation for evaluating autoregressive time series prediction*. Computational Statistics and Data Analysis. 2018; 120:70-83.
- 7- Borup D, Schütte EC. *In search of a job: Forecasting employment growth using Google Trends*. Journal of Business & Economic Statistics. 2022 Jan 2;40(1):186-200.
- 8- Bradley E, Trevor H. *Computer age statistical inference: Algorithms evidence, and data science*. Cambridge University Press; 2021.
- 9- Breiman L, Friedman JH, Stone CJ, Olshen RA. *Classification and regression trees*. CRC press; 1984.
- 10- Bulut L. *Google Trends and the forecasting performance of exchange rate models*. Journal of Forecasting. 2018;37(3):303-315.
- 11- Chandra DP, Manuaba IB. *Prediction of Stock Value Web-scraped Google Trends and Twitter data-driven model*. Procedia Computer Science. 2023 Jan 1; 227:253-61.
- 12- Chen X. *Large sample sieve estimation of semi-nonparametric models*. Handbook of Econometrics. 2007; 6:5549-5632.

- 13- Chu CK, Marron JS. *Comparison of two bandwidth selectors with dependent errors. Annals of Statistics.* 1991;19(4):1906-1918.
- 14- Friedman JH, Hastie T, Tibshirani R. *Applications of the lasso and grouped lasso to the estimation of sparse graphical models.* Technical report. 2010;1-22.
- 15- Gogas P, Papadimitriou T. *Machine learning in economics and finance.* Computational Economics. 2021 Jan;57:1-4.
- 16- Granger CWJ, Pesaran MH. *Economic and statistical measures of forecast accuracy.* Journal of Forecasting. 2000;19(7):537-560.
- 17- Hastie T, Tibshirani R, Friedman JH. *The elements of statistical learning: data mining, inference, and prediction (Vol. 2).* New York: Springer; 2009.
- 18- He Y, Chen C, Li B, Zhang Z. *Prediction of near-surface air temperature in glacier regions using ERA5 data and the random forest regression method.* Remote Sensing Applications: Society and Environment. 2022;100824.
- 19- James WF, Stein CM. *Estimation with quadratic loss.* Proceedings Fourth Berkeley Symposium on Mathematical Statistics and Probability. 1961.
- 20- Khan MA, Abbas K, Su'ud MM, Salameh AA, Alam MM, Aman N, Mehreen M, Jan A, Hashim NA, Aziz RC. *Application of machine learning algorithms for sustainable business management based on macro-economic data: supervised learning techniques approach.* Sustainability. 2022 Aug 12;14(16):9964.
- 21- Liu H, Liu Y, Wang C, Song Y, Jiang W, Li C, Zhang S, Hong B. *Natural Gas Demand Forecasting Model Based on LASSO and Polynomial Models and Its Application: A Case Study of China.* Energies. 2023 May 23;16(11):4268.
- 22- Mansour Y, Ismail R, Alakkari K. *Nowcasting As An Approach To Mitigate Economic Uncertainty In Syria.* Tishreen University Journal- Economic and Legal Sciences Series. 2021;42(6):1-20.
- 23- Mullainathan S, Spiess J. *Machine learning: an applied econometric approach.* Journal of Economic Perspectives. 2017 May 1;31(2):87-106.
- 24- Palimkar P, Shaw R, Ghosh A. *Machine learning technique to prognosis diabetes disease: Random Forest classifier approach.* In: Advanced Computing and Intelligent Technologies: Proceedings of ICACIT 2021; 2022:219-244.
- 25- Paruchuri, H. (2021). Conceptualization of machine learning in economic forecasting. *Asian Business Review*, 11(1), 51-58.
- 26- Persio D, Honchar O. *Recurrent neural networks approach to the financial forecast of Google assets.* International journal of Mathematics and Computers in simulation. 2017; 11:7-13.
- 27- Ranstam J, Cook A. *LASSO regression.* Journal of British Surgery. 2018;105(10):1348-1348.
- 28- Reid SD, Tibshirani RJ, Friedman JH. *A study of error variance estimation in lasso regression.* Statistica Sinica. 2016;35-67.
- 29- Song H, Choi H. *Forecasting stock market indices using the recurrent neural network based hybrid models: CNN-LSTM, GRU-CNN, and ensemble models.* Applied Sciences. 2023 Apr 6;13(7):4644.
- 30- Wald A. *Statistical decision functions.* The Annals of Mathematical. 1949;165-205.