

دراسة مقارنة بين نموذج الانحدار ونموذج بوكس جينكز ونماذج الشبكات العصبية لإنتاج التمور في العراق

م. ماثل كامل ثامر
كلية الإدارة والاقتصاد
جامعة الأنبار

mathilthamir@yahoo.com

المستخلص:

ان اسلوب الشبكات العصبية الاصطناعية يعتبر من الوسائل المهمة الجديدة في بناء النماذج والتحليل وتقييم البيانات لا يعتمد نموذج أو طريقة إحصائية شائعة لتشخيص مسبب لسلوك الظاهرة حيث تتم المعالجة للوصول إلى أفضل نموذج يمثل الظاهرة وبأقل حد ممكن من الاخطاء ويمثل بشكل قريب من الواقع ويمكن استخدامه في أغلب المجالات ومن اهداف تحليل السلاسل بناء نموذج لتفسير سلوكها واستخدام النتائج لغرض التنبؤ بسلوك السلسلة في المستقبل. يهدف البحث الى المقارنة ما بين بعض نماذج الانحدار ونموذج بوكس جينكز ونماذج الشبكات العصبية لاختيار أفضلها لمتغير انتاج التمور في العراق. لقد تم استخدام برامج التحليل الاحصائي SPSS 24 و minitab17 و Eviews9 لتحليل بيانات انتاج التمور للفترة من ١٩٦٣ الى ٢٠١٨ وقد برهنت النتائج كفاءة اسلوب الشبكات العصبية في معالجة النماذج الغير خطية وأنها تعتبر من النماذج الحصينة. الكلمات المفتاحية: انتاج التمور، نموذج الانحدار الخطي، نموذج بوكس جينكيز، نموذج الانحدار الذاتي والاوساط المتحركة، نموذج الشبكة العصبية.

Comparative Study Between Regression Model, Box Jenkins Model and Neural Networks Models for Date Production in Iraq

Lecturer: Mathil Kamil Thamer
College of Administration and Economics
University of Anbar

Abstract:

The method of artificial neural networks is one of the important new methods in building models, analysis and evaluation of data does not adopt a model or a common statistical method to diagnose the behavior of the phenomenon. Where processing is done to reach the best model represents the phenomenon with the least possible errors and represents close to reality and can be used in most areas one of the aims of chain analysis is to build a model to explain its behavior and use the results to predict future behavior of the chain. The aim of this research is to compare some regression models, the Box Jenkins model and the neural networks models to choose the best for the dates production variable in Iraq. Statistical analysis software (SPSS24, minitab17, Eviews9) were used to analyze date production data for the period 1963 to 2018 and the results proved the efficiency of the neural networks method in the treatment of nonlinear models and it is considered to be a strong model.

Keywords: Dates Production, Linear Regression Model, Box Jenkins model, Model Self-regression and Moving Averages, Neural networks model.

المقدمة

ان السلسلة الزمنية تمثل مجموعة من المشاهدات العشوائية لمتغير عشوائي المنشأ من علاقتها بالزمن وفق القوانين الاحتمالية والاحصائية والرياضية. أما تطبيقاً فإن السلسلة الزمنية عبارة عن مجموعة من المشاهدات الخاصة بظاهرة مدروسة وفق فترات زمنية متتابعة وفي الغالب تكون منتظمة، ان التنبؤ باستخدام السلاسل الزمنية ومنها منهجية بوكس-جينكز يعتبر الاسلوب التقليدي في بناء نماذج التنبؤ وكذلك يتم التنبؤ من خلال بناء الشبكات العصبية حيث تستخدم الشبكات العصبية المتعددة الطبقات لغرض التنبؤ. إن من مميزات الشبكات العصبية في هذه الحالة هو ان الشبكة العصبية يمكنها تمثيل المنظومات اللاخطية فضلاً عن امكانيتها على تقليل عدم استقراره البيانات ومعالجة الضوضاء المرافقة الى بيانات السلسلة الزمنية وكذلك القيم الشاذة في السلوك. وعلى اساس هذا المبدأ تم بناء نماذج السلاسل الزمنية من خلال اسلوب بوكس جينكز وبعض النماذج المرتبطة بها، وتم ايضا بناء نموذج لشبكة عصبية بعدد من الطبقات لمتغير انتاج التمر في العراق وفق اسلوب (Perceptrons) اكتسبت الشبكات العصبية في السنوات الاخيرة اهمية متزايدة في معالجة وتحليل السلاسل الزمنية نظراً لما تمتاز به من مرونة كبيرة مقارنة بالطرائق التقليدية المعروفة والمعتمدة في هذا المجال فضلاً عن قدرتها على التعلم والتكيف ذاتياً مع أي نموذج.

مشكلة البحث: ان انتاج التمر في العراق يمثل الدعامة الاساسية للإنتاج الزراعي في العراق ويعتبر التخطيط لها والتنبؤ بكمية الانتاج على الصعيد المحلي او العالمي من الاهمية لهذه الصناعة وذلك لما يحققه من عائد مبني على تحويل منتج زراعي اساسي الى مواد ذات قيمة تجارية وعائد مالي مرتفع لهذا فان مشكلة البحث تنحصر في مدى امكانية بناء نماذج كفوة تساعد المخطط الاقتصادي في تنمية وتطوير زراعة وصناعة انتاج التمر في العراق وتعظيم العائد المالي منها.

أهمية البحث: ان الغرض يتمثل في بناء نماذج متعددة وفق اساليب مختلفة منها نماذج بوكس جينكز ونموذج الشبكات العصبية والمقارنة فيما بينها واختيار أفضلها، ان أي نموذج او سلسلة زمنية خطية كانت او غير خطية سواء كان الخطأ يتبع التوزيع الطبيعي او أي توزيع غير معلوم اخر. فضلاً عن اعتماد منهجية حديثة ذات قدرات ذكية للمعالجة وهي الشبكات العصبية وتقديم معيار كمي رقمي الى الادارة العليا لأجل اتخاذ قرار فعال من خلال استخدام النموذج للتخطيط الاقتصادي.

هدف البحث: بناء نماذج السلاسل الزمنية لبيانات انتاج التمر في العراق وفق بيانات الواقع الفعلي ومعرفة سلوكية هذه السلاسل من ناحية مكونات السلسلة الزمنية كالاتجاه العام والموسمية وغيرها من المركبات وفي مختلف الحالات، وهو يمثل الهدف الاساسي مع المقارنة ما بين عدة طرق واساليب التنبؤ ومنها اسلوب الشبكات العصبية.

الجانب النظري: سوف يتم استعراض اهم نماذج التحليل الاحصائي فيما يخص التنبؤ والتي تتضمن نماذج الانحدار الخطي ونماذج بوكس-جينكز و ثم نماذج الشبكات العصبية مع امكانية تطبيق احداها على بيانات الدراسة.

١. **نماذج الانحدار الخطي:** ان انموذج الانحدار الخطي المتعدد يدرس تأثير مجموعة من المتغيرات المستقلة على المتغير المعتمد ويتم تقدير انموذج الانحدار الخطي المتعدد بطريقة المربعات الصغرى (OLS) لغرض توفيق أفضل معادلة انحدار خطية متعددة المتغيرات مع الاختبارات

الاحصائية المرافقة لها، وان خط الانحدار يهدف الى تصغير مجموع مربعات الانحرافات عن الخط المستقيم الى أدنى حد ممكن، وفق الصيغة الاحصائية الاتية للنموذج: (J. Scott, 2012)، (C.Tofallis, 2009)

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki} + \varepsilon_i \text{-----} (1)$$

$$i = 1.2.3. \dots \dots \dots n$$

ومنه يمكن ان نحصل على نماذج متعددة منها النموذج الخطي البسيط بالصيغة التالية:

$$y_i = \beta_0 + \beta_1 x_{1i} + \varepsilon_i \text{.....} (2)$$

والنموذج التربيعي من الدرجة الثانية بالصيغة التالية:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{1i}^2 + \varepsilon_i \text{.....} (3)$$

ومن الملاحظ على النماذج الاخيرة انها تتعامل مع متغير توضيحي واحد x_{1i} منتج منه متغيرات اخرى من خلال اجراء بعض التحويلات، وتعتمد عملية تقدير النموذج على اسلوب المصفوفات ووفق العلاقة التالية:

$$\hat{\beta} = (x^T x)^{-1} x^T y \text{} (4)$$

حيث ان

$\hat{\beta}$ متجه لمعالم النموذج المقدر بأبعاد $px1$ ، x مصفوفة المتغيرات المستقلة عمودها الاول مساوي للواحد وتكون بأبعاد $n \times p$ ، y مصفوفة المتغير المعتمد وتكون بأبعاد $n \times 1$.

٢. **النموذج المختلط $ARMA(p,q)$:** اما نموذج الانحدار الذاتي والاوساط المتحركة $ARMA(p,q)$ فيوصف بدلالة النماذج السابقة كالآتي: (E.P and Jenkins, et al., 2009)، (Robert H. Shumway, 2017)

$$\phi(L)(y_i) = \theta(L) e_i \text{} (5)$$

ان الذي يجمع في سلوكه مواصفات نماذج الانحدار الذاتي والمتوسطات المتحركة ومن النموذج نجد ان $\phi(L)$ ، $\theta(L)$ تمثل متعددة الحدود لنموذج الانحدار الذاتي ونموذج المتوسطات المتحركة في معالم النموذج، يمكن نحصل من النموذج المختلط على نموذج الانحدار الذاتي $ARMA(p,0)$ اي ان متغير السلسلة الزمنية y_i يعتمد على القيم السابقة للمتغير ويعبر عنه بالصيغة التالية:

$$y_i = \phi_0 y_{i-1} + \phi_1 y_{i-2} + \dots + \phi_p y_{i-p} + e_i \text{.....} (6)$$

٣. **نموذج الشبكة العصبية:** ان الشبكات العصبية الاصطناعية هي انظمة خلوية فيزيائية التي يمكن فيها من كسب وخزن واستعمال المعرفة التجريبية. ومن التعاريف الاخرى هي عبارته عن نماذج رياضية تهتم بمعالجة البيانات والتعلم منها وهي تحاكي بذلك الخصائص التي تتصف بها الانظمة البيولوجية. ومن كل ذلك يمكن ان نصف الشبكات العصبية بانها نظام معالجة بيانات مبني على النماذج الاحصائية والرياضية له مميزات بنقل الاشارة يشابه الشبكة العصبية البيولوجية عند الانسان من خلال مجموعة من النماذج غير الخطية.

بدأت فكرة نشوء الشبكات العصبية الاصطناعية من خلال محاكاة بعض خصائص الشبكات العصبية الحيوية (البيولوجية)، حيث تمتلك الشبكات العصبية الاصطناعية القدرة على

تحليل البيانات وبالتالي إمكانية إجراء التنبؤات لتلك البيانات، أو عزلها وتبويبها إلى أشكال وأنماط مختلفة ومن دون الحاجة إلى أي نموذج مسبق يصف طبيعة العلاقات بين تلك البيانات وما يمكن أن تتضمنه هذه البيانات، كما هو الحال في الطرائق الاعتيادية المعروفة والمستخدمه في حساب التنبؤات.

وبصيغة أكثر حداثة، فقد درست الشبكات العصبية الاصطناعية كأسلوب مقابل إلى أساليب النماذج غير الخطية، بمعنى أن التحليل يعتمد على البيانات المتوفرة مع قليل من العلاقات الأولية حول العلاقات بين المتغيرات وحول النماذج، إذ أن عملية بناء أو إنشاء العلاقات بين متغيرات الإدخال والإخراج تُعنون بخوارزميات التعلم متعدد الأغراض certain-general-purpose 'Learning' algorithms، بمعنى إن الاختلافات بين نماذج الشبكات العصبية الاصطناعية جاءت من الاختلاف في النظريات التي تفسر كيفية تخزين المعلومات والتعلم الذي تقوم به الشبكة العصبية البيولوجية، ومن أبرز هذه النظريات نظرية (Donald Hebb) عام (1949)، إذ تنص هذه النظرية على أن عملية التعلم والذاكرة (Learning & Memory) هي ظاهرة تنتج من قوة نقاط الاشتباك العصبي التي تربط بين العقد العصبية الفعالة، بمعنى آخر إن التحفيز المتكرر (repeated stimulation) لبعض العقد الفعالة في الشبكة العصبية يؤدي إلى حدوث زيادة في قوة الارتباط العصبي بين هذه العقد، وهذه العملية تؤدي إلى تخزين (Encoding) التغيرات في قوة الارتباط العصبي بين العقد الفعالة في الشبكة لاختيار أفضل أوزان وهذا هو معنى التعلم في الشبكة العصبية (Charu C. Aggarwal, 2018) أ. مزايا أسلوب الشبكات العصبية: يمتاز الأسلوب عن غيره من البدائل الإحصائية في عملية تحليل البيانات بالمزايا والعيوب التالية:

- بمجرد أن تقوم الشبكة بتلقي البيانات فإن أدائها لا يتأثر كثيراً عند إدخال أي مجموعة إضافية من البيانات الجديدة والتي لم تكن موجودة في البداية.
- لا تتطلب وجود افتراضات قوية بشأن نوعية البيانات والعلاقة بين المتغيرات كما في الأساليب الإحصائية الأخرى، كما تتطلب توافر عدد كبير من البيانات التجريبية.
- ليس هناك حاجة لترتيب المتغيرات حسب أهميتها، وذلك إذ أن الشبكة تضع تلقائياً أوزان خاصة بها حسب طبيعة البيانات المدخلة إليها
- تعطى قرارات واضحة في النماذج المعروفة، كما يمكنها معالجة مشكلات اتخاذ القرارات والتي تنصف بمدخلات غير واضحة وعشوائية.

ويعاب عليها أنها لا تتعامل مع المشاكل ذات المتغيرات التي لا يمكن صياغتها كمياً، كما تعطي نتائج في بعض الحالات قد لا نستطيع تقديم تفسيراً منطقياً لها. (Charu C. Aggarwal, 2018) ب. **الخلية العصبية الحيوية:** يتميز جسم الإنسان بقدرة عالية على استقبال ومعالجة وإرسال الإشارات من خلال المسارات العصبية في الجهاز العصبي، فالخلية العصبية العضوية تتألف من ثلاثة أقسام حيث تتلقى الخلية العصبية الإشارات من الخلايا الأخرى، وتلك الإشارات تعدل من قيمة الإشارات القادمة ثم تجمع النواة الإشارات القادمة، أي يكون الجهد الناتج عن جمع تلك الإشارات أكبر من حد معين فتسمى عتبة التفعيل أو الحالة الحدية ويتم نقل الإشارة عبر النهايات العصبية ثم تنتشر إلى الخلايا العصبية الأخرى المتصلة، وأن الخلية العصبية الحيوية تمثل البناء الأساسي للشبكة العصبية الاصطناعية، أن الخلية العصبية البيولوجية تتكون من جسم الخلية (Cell Body) وهو

جسم غير منتظم يحتوي على النواة للخلية، التفرعات أو المداخل (Dendrites) وهي عبارة عن اذرع تربط الخلية العصبية بالخلايا المتجاورة وتنقل الإشارات الداخلة الى كل خلية مع الخلايا الأخرى المرتبطة بها، المحور العصبي أو المخرج (Axon) وهو عبارة عن ذراع حيث يقوم بنقل الإشارة الخارجة من الخلية الى الخلايا الأخرى ولذلك كانت هناك محاولات للوصول الى اقرب نقطة لكفاءة الإنسان من خلال تقليد ولو بشكل بسيط هيكله الدماغ البشري باستخدام الشبكات العصبية الاصطناعية (Artificial Neural Networks) والتي تقوم بحل المشاكل دون الحاجة الى وجود خوارزمية حل مما يؤدي الى تقليل كمية البرمجيات وبالتالي تقليل الوقت والكلفة. اما الشبكة العصبية الاصطناعية (Artificial Neural Network)، تُعرف الشبكة العصبية الاصطناعية (ANN) على أنها نظام حسابي مكون من عدد من وحدات المعالجة (Processing Elements) المترابطة مع بعضها وتتصف بطبيعتها الديناميكية والمتوازية في معالجة البيانات الداخلة اليها، أن هدف عمل الشبكات العصبية الاصطناعية هو للوصول الى نموذج لهذه البيانات لغرض التحليل أو التصنيف أو التنبؤ أو أي معالجة اخرى دون الحاجة الى نموذج مقترح لهذه البيانات وبذلك أن الشبكات العصبية الاصطناعية هي تقنيات حسابية مصممة لمحاكاة الطريقة التي يؤدي بها الدماغ البشري مهمة معينة وذلك عن طريق معالجة ضخمة موزعة على التوازي ومكونة من وحدات معالجة بسيطة، هذه الوحدات ما هي الا عناصر حسابية تسمى عصبونات أو عقد (Nodes, Neurons) والتي لها خاصية عصبية، والتي تقوم بتخزين المعرفة العملية والمعلومات التجريبية لتجعلها متاحة للمستخدم وذلك عن طريق ضبط الأوزان. (Yoav Goldberg, Graeme Hirst, 2017)

ج. دالة المجموع **Sum function**: تحدد طريقة وصيغة ادخال المعلومات الى الشبكة العصبية والتي تعرف بالمدخلات وغالباً ما تكون عبارة عن تركيبة خطية (Linear combination) بدلالة الأوزان وتوصف:

$$input = b + \sum_{j=1}^R w_j y_j \dots \dots (7)$$

حيث أن b الحد المطلق (bias)، w_j الأوزان المرتبطة بالمدخلات وهي تقابل المعلمات في نموذج الانحدار، y_j متغيرات الإدخال. (Yoav Goldberg, Graeme Hirst, 2017)

د. دوال الفاعلية **Activation Functions**: تعرف بدالة التحويل (Transfer Function) لأنها تقوم بتحويل المدخلات عبر تفاعلها مع الأوزان من صيغة الى اخرى وتكون منها الخطية وغير الخطية ومنها الدوال الشائعة الاستخدام هي: (Charu C. Aggarwal, 2018)

١. الدالة اللوجستية **Log-Sigmoid Function**: تعد هذه الدالة الأكثر استخداماً عند عقد الطبقة المخفية للشبكة التي مدخلاتها ذات قيم حقيقية بين $(-\infty, \infty)$ أما مخرجاتها لكل عقدة تكون بين $(0,1)$ والصيغة الرياضية لها هي:

$$f(x) = \frac{1}{1 + \exp(-x)} \dots \dots (8)$$

ومشتقتها

$$f'(x) = f(x)[1 - f(x)] \dots \dots (9)$$

٢. الدالة اللوجستية ثنائية القطب **Hyper Bolic Tangent Function**: وتكتب بالشكل الآتي:
 $f(x) = \tanh(x) \dots\dots(10)$

ومشتقتها هي

$$f' = \text{sech}^2(x) = 1 - \tanh^2(x) = (1 - f^2(x)) \dots\dots(11)$$

٣. دالة الجيب أو الجيب تمام **The Sine or Cosine Function**: وتكتب بالصيغة الآتية:
 $f(x) = \cos(x) \text{ or } f(x) = \sin(x) \dots\dots(12)$

ومشتقتها هي

$$f'(x) = -\sin(x) \text{ or } f'(x) = \cos(x) \dots\dots(13)$$

٤. الدالة الثنائية (binary function) وتأخذ الصيغة الاحصائية التالية:

$$f(x) = \begin{cases} 1 & \text{if } x > 0 \\ -1 & \text{if } x < 0 \end{cases}$$

ويمكن توحيد دالة التنشيط المستعملة لعقد الطبقة المخفية في الشبكة أو استعمال دوال مختلفة لكل عقدة فلا يوجد تحديد لنوعية هذه الدوال ولكن أكثر الدوال استعمالاً في عقد الطبقة المخفية هي الدالة اللوجستية.

أما في عقدة طبقة الإخراج فإنه لا يوجد تحديد لدالة التنشيط المستعملة فقد تكون الدالة خطية أو غير خطية فإذا تضمن التدريب على التنبؤ فإن الدالة تكون خطية أما إذا تضمنت عملية التدريب على تصنيف البيانات فإن دالة الإخراج تكون غير خطية.

٥. **تهيئة البيانات**: هي عملية تهيئة البيانات قبل معالجتها (Pre-Processing) لاستخدامها في عملية تدريب الشبكة العصبية لذلك فإن مجال هذه البيانات لابد أن يكون ضمن نطاق دالة التنشيط الحسابية، وتوجد عدة طرق للتهيئة هي: (Yoav Goldberg, Graeme Hirst, 2017)

١. **الصيغة المبسطة Simple Formula**: تعتمد هذه الصيغة على أكبر قيمة في البيانات وتأخذ الصيغة التالية:

$$X_{new} = \frac{x}{x_{max}} \dots\dots(14)$$

٢. **الصيغة الطبيعية Normalized Formula**: تستخدم هذه الصيغة عندما تكون دالة التنشيط في عقدة الإخراج هي الدالة اللوجستية وتعتمد على أقل وأكبر القيم وتوصف بالعلاقة الآتية:

$$X_{new} = \frac{x - x_{min}}{x_{max} - x_{min}} \dots\dots(15)$$

٣. **الصيغة القياسية Standardized Formula**: تستخدم هذه الصيغة عندما تكون دالة التنشيط في عقدة الإخراج هي الدالة الخطية وحسب الصيغة الآتية:

$$X_{new} = \frac{x - \bar{x}}{s} \dots\dots(16)$$

وتوجد صيغ أخرى منها الصيغة الطبيعية المعدلة Adjusted Normalize Formula.

٦. **الشبكات العصبية المتعددة الطبقات Multi-layered neural networks**: تتكون هذه الشبكات من طبقة واحدة أو أكثر من الطبقات المخفية (Hidden Layer) التي بدورها تحتوي على عدد من العقد المخفية (Hidden Nodes)، كما أن للإنسان وحدات إدخال توصله بالعالم الخارجي وهي حواسه الخمس، فذلك الشبكات العصبية تحتاج لوحدة إدخال، ووحدات معالجة

يتم فيها عمليات حسابية تضبط الأوزان ونحصل من خلالها على ردة الفعل المناسبة لكل مدخل من المدخلات للشبكة، وتتميز هذه الشبكات بقدرتها على حل مسائل أكثر تعقيداً من المسائل التي يمكن حلها في الشبكات ذات الطبقة المفردة بسبب وجود الطبقات الوسيطة التي تعطي مرونة أكثر في بناء الدوال الناقلة بين المدخلات والمخرجات، وعلى الرغم من أن تدريب هذه الشبكات يستغرق وقتاً أطول لكن التدريب لهذه الشبكات ناجح أكثر من غيره، ونود أن نشير انه عندما تكون وحدات الإدخال متغيرات توضيحية $(x_p, \dots, x_1)$ فإن أنموذج الشبكات العصبية سيكون مشابه الى أنموذج الانحدار المتعدد غير الخطي (Non-Linear Multiple Regression Model) الذي يكتب بالشكل الاتي: (R. Cole, Matt, 2018)

$$y_t = f(x_1, \dots, x_p) + e_t \dots \dots \dots (17)$$

وعندما تكون وحدات الادخال تمثل إزاحات (Lags) المتغير المعتمد وهي $(y_{t-p}, \dots, y_{t-1})$ فإن أنموذج الشبكة سيكون مشابه الى أنموذج الانحدار الذاتي غير الخطي Non-Linear- Autoregressive Model والذي يكتب بالصيغة الآتية:

$$y_t = f(y_{t-1}, \dots, y_{t-p}) + e_t \dots \dots \dots (18)$$

المستوى المخفي Hidden Level: يمثل المستوى الثاني للشبكة العصبية والذي يلي المستوى الأول (مستوى المدخلات) وقد يحتوي على عدد من العقد المخفية (Hidden Nodes) ولكل عقدة وزن يربطها مع المستوى السابق (مستوى المدخلات) ووزن يربطها مع المستوى اللاحق (مستوى المخرجات) كما هو الحال في الشبكات العصبية ذات الطبقة المخفية الواحدة (Single-Layer Network).

مستوى المخرجات Output Level: يمثل المستوى الأخير في الشبكة العصبية الاصطناعية الذي هو عبارة عن مخرجات الشبكة العصبية ويمكن أن يكون للناتج عقدة واحدة أو أكثر، وترتبط هذه المستويات الثلاثة بعدد من العقد (Nodes) والتي تمثل نقاط الارتباط (التشابك) العصبي بين مستويات الشبكة العصبية ويكون عملها بأن تستلم المدخلات من المستوى الأعلى منها وتنتشر المخرجات الى المستوى الأدنى عن طريق الأوزان (Weights) وأن هذه الأوزان توضح مدى قوة الارتباط العصبي بين المستويات إذ أن لكل عقدة مخفية وزن يربطها مع المستوى السابق ووزن يربطها مع المستوى اللاحق، وعليه فإن الشبكة العصبية بشكل عام تحتوي على ثلاثة طبقات من الأوزان وهي: (R. Cole, Matt, 2018)

- طبقة أوزان مستوى المدخلات والمستوى المخفي (Input to Hidden Weights): حيث نرمز لها أما بالرمز (w_{ij}) إذ أن (i) يرمز الى عقد الادخال $(i=1, \dots, p)$ و (j) يمثل عدد العقد المخفية، أو نرمز لها (w_{ih}) إذ أن (h) تمثل العقد المخفية.

- طبقة أوزان بين المستويات المخفية (Between Hidden Weights): وتظهر هذه الأوزان فقط في حالة الشبكات العصبية المتعددة الطبقات إذ أن لكل طبقة مخفية عدد من الأوزان.

- طبقة أوزان المستوى المخفي ومستوى المخرجات (Hidden to Output Weights): ويرمز لها بالرمز (w_j) أو (w_{ho}) إذ أن (o) يمثل الناتج (output).

ز. **نموذج الشبكة العصبية:** وبشكل عام فإن أي أنموذج للشبكات العصبية بطبقة مخفية واحدة وبعدها عقد مخفية (r) وعدد عقد الادخال (p) يمكن أن يكتب بالشكل الاتي: (Yoav Goldberg, Graeme Hirst, 2017)

$$\hat{y}_t = \varphi_0 \{ w_0 + \sum_{j=1}^r w_j * f(w_{1j} + \sum_{i=1}^p w_{ij} y_t) \} \dots\dots\dots (19)$$

إذ أن:

r عدد العقد المخفية، p عدد عقد الإدخال، w_{1j} تمثل الأوزان التي تربط الثابت وبين العقد المخفية، w_0 الوزن الذي يربط الثابت مع الناتج، w_j الأوزان التي تربط بين العقد المخفية والناتج، w_{ij} الأوزان التي تربط بين عقد المدخلات والعقد المخفية، φ_0 دالة التنشيط (التفعيل في مستوى المخرجات)، f دالة التنشيط (التفعيل في مستوى الخفي)، y_t تمثل متغيرات الإدخال. (R. Cole, Matt, 2018)

ح. **خوارزميات التعلم في الشبكات العصبية الاصطناعية:** تهدف خوارزميات التعلم في الشبكات العصبية الاصطناعية إلى تحديد قيمة الأوزان المرجحة من خلال تدريب الشبكة بإحدى الطرق الثلاثة الآتية: (Yoav Goldberg, Graeme Hirst, 2017)

١. **التعلم المراقب (Supervised Learning):** يعتمد التعلم المراقب للشبكات العصبية الاصطناعية على فكرة عرض البيانات التدريبية أمام الشبكة على هيئة زوج من الأنماط يسمى نمط الإدخال (Input Pattern)، ونمط الإخراج (Output Pattern)، وكلاً منهما مرتبط بمتجه مخرجات مستهدف (Target Output Vector)، وتستخدم الشبكة الفرق بين مخرجات الشبكة والمخرجات الحقيقية في حساب دالة الخطأ (δ) (Delta Function)، التي تستخدمها بعد ذلك في تعديل الأوزان لتقليل الفرق بين المخرجات للشبكة والمخرجات الحقيقية، ويتم تعديل الأوزان لايجاد أفضل النتائج وذلك باستخدام دالة تحديث الأوزان أو دالة التعلم (Learning Function).

٢. **التعلم غير المراقب (Unsupervised Learning):** تقوم أساليب التعلم الذاتي للشبكات العصبية الاصطناعية على أساس قدرتها على اكتشاف الملامح المميزة لهيكل المدخلات من خلال مجموعة من البيانات يتم تعميمها على باقي المخرجات، وذلك من دون معرفة مسبقة ومن دون عرض أمثلة لما يجب أن تنتجه، وذلك على عكس المبدأ المتبع في أسلوب التعلم المراقب بمعنى أن للشبكات العصبية القدرة على التنظيم الذاتي لأوزان ترابطاتها.

٣. **التعلم بإعادة التدعيم (Reinforcement Learning):** تقوم عملية التدعيم على أساس الخلط بين طريقة التعلم المراقب والتعلم غير المراقب، حيث لا يفصح للشبكة العصبية عن القيم الحقيقية للمخرجات كما هو الحال في التعلم غير المراقب، ولكن يشار إلى الشبكة بصحة نتائجها أو خطئها كما في طريقة التعلم المراقب.

ط. **خوارزمية الانتشار العكسي للخطأ:** أن منهجية الانتشار العكسي (EBP) للخطأ طورت في عام 1970 من قبل عدة علماء إذ أن الخطوات الأساسية فيها هو حساب خطأ مستوى المخرجات لتحديث أوزان طبقة المستوى المخفي-المخرجات، ثم حساب خطأ مستوى المخفي لتحديث أوزان طبقة مستوى المدخلات-المخفي، وبعد ذلك نحسب مخرجات الشبكة بالأوزان الجديدة لتستمر العملية في حساب الخطأ وتحديث الأوزان للوصول إلى أقل خطأ في الشبكة العصبية.

إن خوارزمية الانتشار العكسي تستخدم في تدريب الشبكات العصبية كاملة الارتباط ومتعددة الطبقات وغير الخطية، ويتم تنفيذ هذه الخوارزمية من خلال ثلاث مراحل وكالاتي: (مرحلة الانتشار الأمامي للخطأ، مرحلة الانتشار الخلفي للخطأ، ثم مرحلة توليف أوزان الشبكة). تبدأ العملية بحساب قيمة الناتج (output) بموجب نموذج الشبكات العصبية ذات الطبقة المخفية

الواحدة بالاعتماد على القيم الابتدائية للأوزان وعلى المدخلات من دون إضافة الخطأ، ثم يتم إيجاد الخطأ بين المخرج المطلوب والمخرج الفعلي وترجع بهذا الخطأ عكسياً من الطبقة الأخيرة إلى الطبقة الخلفية ثم أخيراً إلى طبقة المدخلات، وفي أثناء الرجوع (الارتداد) يتم تغيير الأوزان في الاتجاه الذي يرجع الخطأ إلى النقصان وأحياناً إلى الصفر وهذه مرحلة الانتشار الخلفي للخطأ، وتستعمل هذه الطريقة التدريجية مع الشبكات ذات التغذية الأمامية إذ أن التغذية الأمامية يرجع إلى بنية الشبكة أما انتشار الخطأ العكسي فيرجع إلى طريقة التدريب التي تستعمل مع هذه الشبكة.

وكما هو معلوم أن الشبكة لها ثلاث مستويات وعدد عقد الإدخال هي (i) وعدد العقد في الطبقة المخفية هي (j) وعدد عقد الإخراج هي (k) - تبدأ هذه الخطوة عند طبقة المخرجات وفيها نحدد مجموع الاشارات التي تدخل على كل من خلاياه وكما مبين ادناه:

$$N_k = \sum w_{jk} o_j \dots\dots\dots (20)$$

بعد مرور هذه القيمة على الخلية والممثلة في الدالة التنشيطية اللوجستية يكون المخرج وحسب المعادلة الآتية:

$$O_k = \frac{1}{1 + \exp(-N_k)} = f(N_k) \dots\dots\dots (21)$$

- يتم تحديد مقدار الخطأ العشوائي δ_k الذي يمكن الحصول عليه من اشتقاق متوسط مربعات الخطأ بالنسبة إلى الناتج O_K إذ أن:

$$\delta_k = (t_k - O_k) * f'(N_k) = (t_k - O_k) * O_k(1 - O_k) \dots\dots\dots (22)$$

وبالحصول على هذه القيمة يمكن أن نصح كل الأوزان التي تربط الطبقة الخفية بطبقة المخرجات وبالسرع المطلوبة حسب نسبة التعلم η حسب الصيغة الآتية:

$$w_{jk(new)} = w_{jk(old)} + \eta \delta_k O_j \dots\dots\dots (23)$$

حيث ان δ_k يمثل مقدار الخطأ العشوائي ويعرف:

$$\delta_k = (t_k - O_k) * f'(N_k)$$

ويتم تطبيق هذه الخطوات على كل المدخلات المعدة للتدريب ولمرات عديدة حتى تصل الشبكة إلى أقل نسبة خطأ ممكنة ولتصبح الشبكة جاهزة للاستعمال. (Charu C. Aggarwal, 2018)

ي. مراحل التنبؤ بواسطة الشبكات العصبية الاصطناعية

- تحديد المتغيرات، جمع البيانات، معالجة البيانات.

- تجزئة البيانات: في الخلية العصبية الاصطناعية تجزأ البيانات إلى ثلاث مجاميع لاستعمالها في

التحليل والتنبؤ وكما يلي: (Yoav Goldberg, Graeme Hirst, 2017)

١. **مجموعة التدريب أو التعلم:** وهي مجموعة جزئية من بيانات السلسلة الزمنية إذ تستعمل هذه المجموعة في تقدير الأوزان الأولية بنسبة أولية وأن النسبة التي يتم اختيارها في مجموعة التعلم يجب أن تكون بحجم معين لتعطي مواصفات ومميزات للمتغيرات لتدريب الشبكة والتي تستخدم لتقدير الأوزان وتدريب الشبكة على التنبؤ.

٢. **مجموعة الاختبار:** وهي جزء من بيانات السلسلة الزمنية ولعينة لم تستعمل في التدريب حيث تقوم باختبار قابلية الشبكة على التنبؤ بعد عملية التعليم، بعض البرمجيات تحدد 70% للتعلم و30% للاختبار.

٣. **مجموعة التقويم:** تقوم بمراقبة نهائية لتقويم إداء تدريب الشبكة العصبية بصورة عامة.
- ❖ **تحديد معمارية الشبكة العصبية:** لتحديد هيكلية الشبكة العصبية إذ يتطلب من مصمم الشبكة مهارة عالية وخبرة تجريبية في تحديد الخطوات الأتية: (Yoav Goldberg, Graeme Hirst, 2017)
١. **تحديد عدد عقد الإدخال:** أن المتغيرات في أنموذج السلسلة الزمنية تكون بدلالة المتغيرات المزاحة أي أن:

$$Y_t = f(Y_{t-1}, Y_{t-2}, \dots, Y_{t-p})$$

- إن تحديد هذه المتغيرات تمثل مشكلة كبيرة تواجه مصمم الشبكة العصبية وقد اقترحت عدة طرق منها: **الطريقة التي تعتمد على معاملات الارتباط الذاتي المعنوية للسلسلة الزمنية المستقرة.**
- او **الطريقة التي تعتمد على أسلوب تجميع البيانات** فإذا كانت شهرية فان عدد عقد الإدخال هي: $P = 12$ ، وإذا كانت فصلية فأن $P = 4$.
٢. **تحديد عدد الطبقات المخفية:** في أغلب تطبيقات التنبؤ باستخدام الشبكات العصبية المتعددة الطبقات يفضل استخدام طبقة مخفية واحدة.
٣. **تحديد عدد العقد في الطبقة المخفية:** تواجه مصمم الشبكة العصبية مشكلة أخرى الا وهي تحديد عدد العقد في الطبقة المخفية، أن عدد العقد المخفية للشبكة العصبية يحدد من قبل المستخدم وللبدء بتدريب الشبكة العصبية لا بد من اعطاء تقدير أولي لعدد العقد المخفية. أن بعض الباحثين يختارها بعدد المتغيرات المزاحة زمنياً أي بعدد عقد الإدخال في حين البعض الآخر يعتمد على مقياس AIC, BIC, AIC_C مع وجود طرق أخرى.

❖ **تحديد عقد الإخراج:**

- **اختيار دوال التنشيط:** غالبية الشبكات العصبية الحالية تستخدم الدالة اللوجستية في الطبقة المخفية والدالة الخطية عند طبقة المخرجات.
- **تدريب الشبكة العصبية:** أن عملية تدريب الشبكات العصبية على التنبؤ تتم من خلال تعديل أوزان الشبكة للحصول على أقل خطأ بالاعتماد على خوارزمية الانتشار الخلفي (BP)، وأن الهدف من التدريب هو العثور على مجموعة من الأوزان بين الخلايا العصبية التي تحدد الحد الأدنى لدالة الخطأ ومن أهم خطوات تدريب الشبكة هي: (R. Cole, Matt, 2018)
١. **تحديد قيم الأوزان الأولية:** وتُعد نقطة البداية لعملية التدريب وغالباً ما يتم اختيار قيم عشوائية صغيرة للأوزان والحد المطلق تقع ما بين (1,-1).
٢. **عدد تكرارات التدريب:** توقف عملية تدريب الشبكة العصبية وذلك من خلال تقليل خطأ مجموعة التقييم أي الوصول إلى خطأ مسموح به وبذلك تنتهي عملية التدريب، او ان تتم إيقاف التدريب بعد تقييم عدد محدد مسبقاً من التكرارات.
٣. **معدل التعلم (η) وعامل الزخم (α):** يعتبر معدل التعلم (η) من العوامل المؤثرة على عملية سرعة تحديث الأوزان في الشبكة العصبية ، ويتم تدريب الشبكة باستخدام خوارزمية الانتشار الخلفي التي تعتمد على تقليل مجموع مربعات الأخطاء والمعرفة بالمعادلة الأتية :

$$E = \frac{1}{2} \sum_h^M E_h = \frac{1}{2} \sum_h^M \sum_i^N (t_{hi} - o_{hi})^2 \dots \dots \dots (24)$$

حيث ان: E : الخطأ الكلي لجميع العقد. E_h : يمثل الخطأ للعقدة h ، h : تشير الى مستوى مجموع

المدخلات، t_{hi} : تمثل المخرجات الحقيقية، O_{hi} : تمثل مخرجات الشبكة.
ان حجم الشبكة يؤثر على عملية التعلم حيث ان الشبكة الكبيرة تحتاج الى وقت أكبر للتعلم والوصول الى اقل خطأ، ويتأثر بمعدل التعلم (η) ومعامل الزخم (α) فكلما كانت عملية التعلم بطيئة نعطي قيمة أكبر الى (η) والتي تؤدي الى تغيير الوزن اي تحديثه يكون كبير مع انخفاض التكرارات لعملية التدريب للحصول على خطأ مقبول.

أما بالنسبة الى عامل الزخم (α) يختار بحيث يجعل عملية التعلم ومقدار التغير في الوزن مستقر نسبياً واجمالياً يعمل على تقليل التكرارات للوصول الى اقل خطأ، فالقيمة الكبيرة لها ستؤدي الى زيادة سرعة التعلم وتؤخذ بنظر الاعتبار عندما تكون عملية الوصول الى اقل خطأ بطيئة، ويكون مداه بين الصفر والواحد ($0 \leq \alpha < 1$).

- **التنفيذ:** وهي من أهم المراحل حيث تختبر الشبكة من حيث قدرة التكيف مع حالة التغير في الدورة وامكانية إعادة التدريب والوصول الى أقل مربع خطأ عند تغير البيانات، ويتم بناء الشبكات العصبية عن طريق البرمجة (أي أن الشبكة العصبونية هي عبارة عن برنامج حاسوبي) ويتكون من عدد كبير من الدورات التي يزيد استعمالها لحل مشكلة معينة. (R. Cole, Matt , 2018)

٤. الجانب التطبيقي:

تم الاعتماد على البيانات الخاصة بانتاج التمور في العراق للفترة ١٩٦٣-٢٠١٨ والتي جمعت من البيانات الاحصائية الخاصة بالانتاج الزراعي في الجهاز المركزي للإحصاء وقد وضحت في الجدول (١) وذلك لاهمية النخلة التي تعد رمز العراق.

الجدول (١): انتاج التمور في العراق للفترة ١٩٦٣-٢٠١٨

السنة	المجموع	السنة	المجموع
1963	420000	1991	566218
1964	320000	1992	447837
1965	310000	1993	612580
1966	280000	1994	675820
1967	380000	1995	881020
1968	330000	1996	797450
1969	260000	1997	750140
1970	480000	1998	913020
1971	300000	1999	763720
1972	450000	2000	931540
1973	310000	2001	906790
1974	385000	2002	919470
1975	٥97190	2003	868390
1976	371980	2004	448380
1977	578310	2005	404030

السنة	المجموع	السنة	المجموع
1978	389030	2006	432360
1979	519170	2007	430860
1980	596920	2008	476320
1981	370180	2009	507002
1982	373750	2010	566829
1983	345290	2011	619182
1984	251440	2012	655450
1985	390230	2013	676111
1986	434210	2014	662447
1987	323540	2015	602348
1988	356330	2016	615211
1989	488390	2017	618818
1990	544931	2018	646163

حيث تم الاعتماد على بناء النماذج التالية:

أ. نموذج الانحدار الخطي البسيط من علاقته بالسنوات والذي بين ان هناك تزايد في الانتاج بمقدار معلمة الميل المقدرة والتي بلغت ٤٦٤,٦٦٨١ بالإضافة الى معنوية النموذج والمعالم المقدرة مع ارتفاع مجموع اخطاء النموذج وانخفاض معامل التحديد.

الجدول (٢): يبين نتائج نموذج الانحدار الخطي البسيط

Dependent Variable: DATPRO

Method: Least Squares

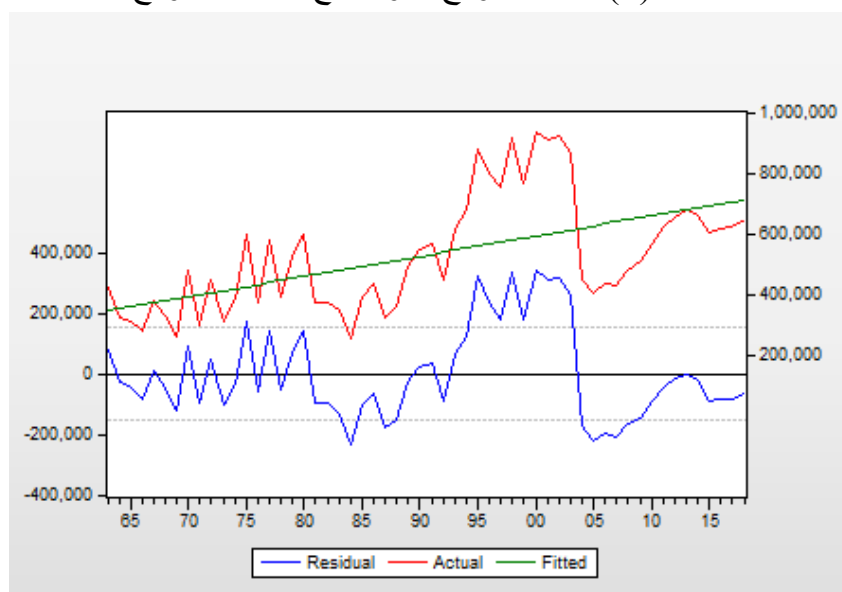
Date: 09/27/19 Time: 19:10

Sample: 1963 2018

Included observations: 56

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-12771751	2548473.	-5.011531	0.0000
YEAR	6681.464	1280.276	5.218769	0.0000
R-squared	0.335266	Mean dependent var	527703.5	
Adjusted R-squared	0.322957	S.D. dependent var	188199.1	
S.E. of regression	154855.2	Akaike info criterion	26.77343	
Sum squared resid	1.29E+12	Schwarz criterion	26.84576	
Log likelihood	-747.6560	Hannan-Quinn criter.	26.80147	
F-statistic	27.23555	Durbin-Watson stat	0.650198	
Prob(F-statistic)	0.000003			

الشكل (١): يبين النموذج الموفق مع اخطاء النموذج



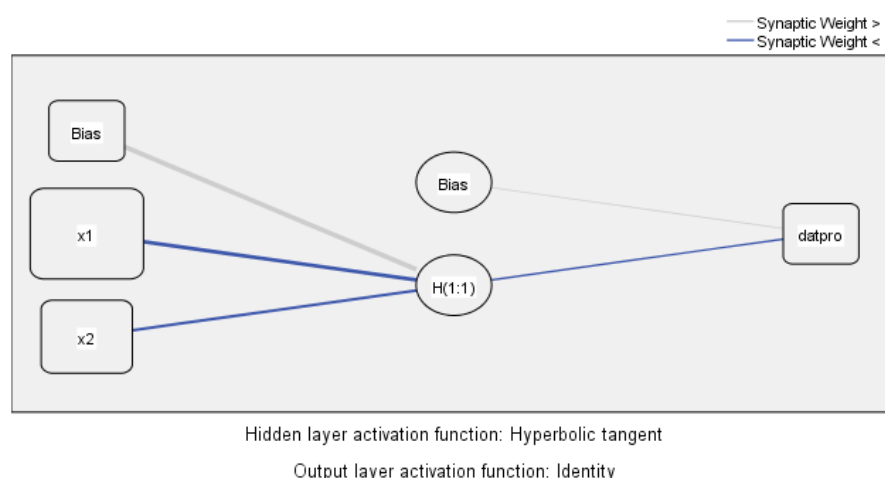
ب. نموذج الانحدار الذاتي والمتوسطات المتحركة حيث تم بناء مجموعة من النماذج بلغ عددها ٢٠ نموذج بعد حساب معاملات الارتباط الذاتي والارتباط الذاتي الجزئي مع اختبار استقرارية السلسلة الزمنية حيث تم تحقيق الاستقرارية في التباين من خلال التحويل اللوغارتمي ثم الفرق الاول لتحقيق الاستقرارية في المتوسط وكان افضل نموذج من حيث المقاييس النموذج $ARIMA(2,1,2)$ والذي وضحت نتائجه في الجدول التالي حيث نلاحظ انخفاض الاخطاء مقارنة بمقارنة بالنموذج الخطي .

الجدول (٣): يبين نتائج نموذج الانحدار الذاتي والمتوسطات المتحركة

Dependent Variable: DLOG(DATPRO)				
Method: ARMA Maximum Likelihood (BFGS)				
Date: 09/27/19 Time: 18:41				
Sample: 1964 2018				
Included observations: 55				
Convergence achieved after 19 iterations				
Coefficient covariance computed using outer product of gradients				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.009638	0.021181	0.455040	0.6511
AR(1)	-1.292471	0.156295	-8.269439	0.0000
AR(2)	-0.526703	0.120226	-4.380955	0.0001
MA(1)	0.882240	0.128949	6.841772	0.0000
SIGMASQ	0.047177	0.009700	4.863392	0.0000
R-squared	0.295391	Mean dependent var		0.007833
Adjusted R-squared	0.239022	S.D. dependent var		0.261141
S.E. of regression	0.227804	Akaike info criterion		-0.019667
Sum squared resid	2.594736	Schwarz criterion		0.162818
Log likelihood	5.540844	Hannan-Quinn criter.		0.050901
F-statistic	5.240335	Durbin-Watson stat		1.866734
Prob(F-statistic)	0.001325			
Inverted AR Roots	-.65-.33i	-.65+.33i		
Inverted MA Roots	-.88			

ج. نموذج الشبكة العصبية: تم بناء نموذج الشبكات العصبية لمتغير كمية انتاج التمور وبمرحلة تدريب بلغت (70%) من البيانات ومرحلة اختبار (30%)، وبطريقة مخفية واحدة مكونة من عقد واحدة

بعد تحويل البيانات الى الصيغة القياسية وبدالة تحويل (hyperbolic tangent) للعقد في الطبقة المخفية ودالة اخراج (identity) لنتائج المخرج النهائي ، وكانت النتائج موصوفة بالشكل التالي لافضل شبكة مقدر.



وبمصفوفة الازان التالية ما بين طبقة الادخال والطبقة المخفية بعقد واحدة:

Parameter Estimates

Predictor		Predicted	
		Hidden Layer 1 H(1:1)	Output Layer datpro
Input Layer	(Bias)	2.138	
	x1	-1.978-	
	x2	-1.481-	
Hidden Layer 1	(Bias)		.668
	H(1:1)		-1.452-

وبمقاييس الكفاءة التالية الخاصة لمرحلة التدريب والاختبار:

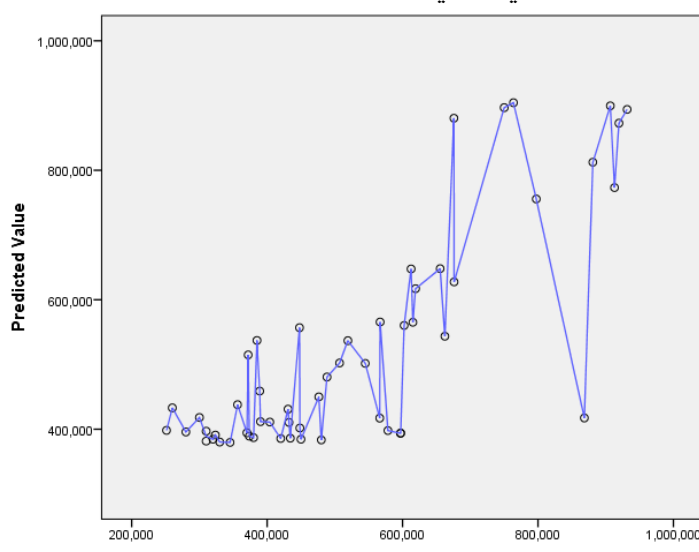
Model Summary

Training	Sum of Squares Error	1.694
	Relative Error	.395
	Stopping Rule Used	1 consecutive step(s) with no decrease in error ^a
	Training Time	0:00:00.00
Testing	Sum of Squares Error	1.057
	Relative Error	.189

Dependent Variable: datpro

a. Error computations are based on the testing sample.

وبالقيم التنبؤية التي زودتنا بالشكل البياني التالي:



د. مقارنة نموذج الانحدار الخطي ونموذج بوكس جينكز مع نموذج الشبكة العصبية: لأغراض المقارنة بين النماذج الثلاثة تم اعتماد المعيار الاحصائي (MSE) لاختيار النموذج المناسب للسلسلة الزمنية لإنتاج التمور، حيث تم اعتماد نموذج الشبكة العصبية لسلسلة انتاج التمور وكما موضح في الجدول الاتي:

الجدول (٤): المقارنة بين نموذج الانحدار الخطي وبوكس جينكز ونموذج الشبكة العصبية

نموذج الشبكة العصبية	نموذج بوكس جينكز	نموذج الانحدار	النموذج
0.0207	0.0519	2.388889e ¹⁰	MSE

من خلال ما تقدم يمكن القول بان اسلوب الشبكات العصبية أكثر ملائمة لتحليل بيانات السلسلة الزمنية من طريقة الانحدار الخطي وطريقة بوكس جينكز.

الاستنتاجات والتوصيات

أولاً. الاستنتاجات: ان نماذج الانحدار غير كفوة كما ان بوكس جينكز تعاني من مشاكل متعددة منها عدم امكانية تحديد رتبة النموذج سواء كان الى نموذج الانحدار الذاتي او الى المتوسطات المتحركة لعدم مطابقة واقع البيانات مع السلوك النظري لهذه النماذج وهذا ما تم معاناته عند بناء النموذج. كما برهنت النتائج كفاءة اسلوب الشبكات العصبية الاصطناعية في معالجة النماذج غير الخطية، ان نماذج الشبكات العصبية الاصطناعية في الأغلب هي نماذج حصينة لأنها تتدرب على البيانات.

ثانياً. التوصيات: يمكن أيجاز اهم التوصيات التي تم التوصل اليها في ضوء الاستنتاجات، بما يلي: اعتماد أسلوب الشبكات العصبية الاصطناعية في البحوث والدراسات ذات الصلة بالسلاسل الزمنية. مع اعتماد الشبكات العصبية الاصطناعية في معالجة نماذج السلاسل الزمنية ذات الانماط المختلفة والمتنوعة والتي لا يتوزع فيها الخطأ وفقاً للتوزيع الطبيعي والتوصية في اعتماد الشبكات العصبية الاصطناعية في معالجة السلاسل الزمنية متعددة المتغيرات "Multivariate" ونماذج دالة التحويل الخطية.

المصادر:

1. C.Tofallis, (2009), Least Squares Percentage Regression. Journal of Modern Applied Statistical Methods. 7: 526-534.
2. Armstrong Analysis ،2-J. Scott, (2012), Illusions in Regression. International Journal of Forecasting (forthcoming). 28 (3): 689.
3. Box, George، 3- E.P and Jenkins, GwilymM .and Reinsel ,Gregory C., (2013), Time Series Analysis for casting and Control
4. Robert H. Shumway, (2017), Time Series Analysis and Its Applications: With R Examples (Springer Texts in Statistics), usa
5. Charu C. Aggarwal, (2018), Neural Networks and Deep Learning: A Textbook 1st ed. Edition, spinger, usa.
6. Yoav Goldberg , Graeme Hirst, (2017), Neural Network Methods in Natural Language Processing (Synthesis Lectures on Human Language Technologies).
7. R. Cole, Matt, (2018), Hands-On Neural Network Programming with C#: Add powerful neural network capabilities to your C# enterprise applications, packt, USA.