جامعة المنصورة
كلية التجارة
قسم الإحصاء التطبيقي و التأمين

مقارنة للتنبؤ بالسلاسل الزمنية متعددة المتغيرات باستخدام الشبكات العصبية الاصطناعية ونموذج دالة التحويل

(دراسة تطبيقية)

الأستاذة الدكتورة / فاطمة علي محمد عبد العاطي
أستاذ الاحصاء التطبيقي والتأمين
كلية التجارة – جامعة المنصورة

الأستاذ الدكتور / محمد توفيق اسماعيل البلقيني
أستاذ الرياضيات و الاحصاء الاكتواري
كلية التجارة – جامعة المنصورة

بحث مقدمة من

سعاد محمد سالم عبد المعطي

كجزء من متطلبات الحصول علي درجة الدكتوراه في

الإحصاء التطبيقي

٢٠١٨

المستخلص

١. تناولت هذه الدراسة مقارنة دقة التنبؤ بالسلاسل الزمنية متعددة المتغيرات باستخدام نماذج دالة التحويل ونماذج الشبكات العصبية الاصطناعية بالتطبيق علي بيانات الانتاج والاستهلاك من الغاز الطبيعي في الفترة من " ١٩٩٠ - ٢٠١٥ " .

٢. تتمثل أهمية هذه الدراسة في معرفة النموذج أكثر دقة للتنبؤ بالسلاسل الزمنية متعددة المتغيرات بعد مقارنة نماذج دالة التحويل ونماذج الشبكات العصبية .

٣. تهدف هذه الدراسة إلي مقارنة الاساليب الذكية و الحديثة " الشبكات العصبية الاصطناعية - دالة التحويل " إلي جانب الاساليب التقليدية ومدارسة الفجوة بين الانتاج و الاستهلاك من الغاز الطبيعي .

٤. وقد تعددت الاساليب الاحصائية إلا أن المشكلة تكمن في استخدام الاسلوب الأنسب لطبيعة البيانات و قد توصلت النتائج إلي تميز الشبكات العصبية و دالة التحويل علي الاساليب التقليدية و إن كانت الأفضلية لدالة التحويل علي الشبكات العصبية الاصطناعية خاصة مع السلاسل الزمنية طويلة الأجل .

(١ – ١) مقدمة

- لقد أصبح الاتجاه العام في البحوث والدراسات الاقتصادية والادارية هو استخدام طرق القياس الكمي ووسائل الاقناع الاحصالية وذلك لتحديد الخصائص و إبراز الاتجاهات العامة للظواهر الاقتصادية والاجتماعية و الادارية وتحليل العلاقات المتشابكة و المتبادلة بين الظواهر علي أساس موضوعي غير متحيز . وعلم الاحصاء يعطي العديد من الطرق و الاساليب اللازمة للقيام بالدراسات و البحوث علي أساس من القياس لحركة العديد من المتغيرات المحددة للظواهر موضوع الدراسة . حيث أنه توجد علاقة ترابط وثيقة بين التخطيط العلمي و المستقبل العلمي و التوقع والتخطيط المدروس أصبح يشكل جزءا لا يتجزأ من علم المستقبليات ، فالتخطيط هو التنبؤ بما سيكون عليه المستقبل و الاستعداد له ، ولتحقيق أهداف معينة في إطار زمني محدد من خلال خطط يتم إعدادها بإتقان لتحقيق أهداف محددة ، و كأن الترابط بين التخطيط العلمي و التنبؤ وثيقا جدا حيث أن التنبؤ يتعلق بعلم المستقبليات لأن هذا العلم يعتمد علي التنبؤ و التوقع العلمي اعتمادا كبيرا حيث يتم من خلاله التعرف علي احتمالات المستقبل الايجابية و السلبية و الاستعداد لمواجهتها من جانب وتوسيع دائرة الاحتمالات الايجابية ليأتي دور التخطيط العلمي مستثمرا لعملية التنبؤ بما ينعكس ايجابيا علي المجتمع والمستثمر علي وجه الخصوص .

- **(١ – ٢) مشكلة الدراسة:**

- تشهد مصر نقص في كميات الغاز الطبيعي وليست هذه الأزمة وليدة اللحظة لكن نتيجة تراكمات عديدة تصل لدرجة الاختناق في المحافظات المصرية أو حسب الانتاج المتوقع من ذلك الحقل أم من حقول آخري مماثلة في المستقبل القريب .
- – تتمثل مشكلة البحث في عملية بناء نموذج احصائي للتنبؤ بكمية الغاز الطبيعي في مصر لما تشهده مصر من نقص خلال ٣٤ عاما القادمين . ولكن في الآونة الأخيرة اكتشف حقل غاز يزيد من انتاج الغاز الطبيعي في مصر بنسبة غير معرفة حتي الآن
- وبالتالي قد تنخفض تلك السنوات أو قد لا نحتاج الغاز لفترات طويلة وجاري توضيح ذلك .

(١-٣) أهداف الدراسة :

يمكن بلورة أهداف البحث في النقاط التالية :

-إلقاء الضوء علي توضيح طبيعة عمل اسلوب الشبكات العصبية الاصطناعية (ANN) Artificial Neural Networks and its Architecture توضيح أهمية اختبار مدي كفاءتها في عملية التنبؤ .

- يهدف البحث لدراسة تحليل بيانات سنوية للناتج المحلي الاجمالي وأهم المتغيرات المؤثرة عليها باستخدام الشبكات العصبية ودالة التحويل مقارنة بالأساليب التقليدية .
- يهدف البحث لدراسة تحليل بيانات سنوية للاستثمار الأجنبي المباشر و أهم المتغيرات المؤثرة عليها
- يهدف لدراسة المقارنة بين الشبكات العصبية الاصطناعية والدالة التحويلية ودراسة الأفضلية بينهم باعتبارهما من الأساليب الإحصائية الحديثة

(١ – ٤) أهمية البحث:

- التنبؤ بإنتاج الغاز الطبيعي في مصر باستخدام اسلوب Box –Jenkins و الشبكات العصبية Neural Networks وذلك لمساعدة الجهات المسئولة عن الإنتاج أن تكون علي الوضع الصحيح من انتاج مواقعها المختلفة من أجل المساعدة في وضع الخطط و الدراسات المستقبلية سواء في المواقع البرية أو المواقع البحرية .
- التنبؤ بإنتاج المواقع المختلفة للجهات المسئولة وذلك لإثبات مدي صحة اتجاه استراتيجية الجهة من عدمه .
- حاجة المستثمرون لمعرفة معايير أو مؤشرات موضوعية تساهم في اتخاذ القرار نحو الاستثمار في مجال الغاز الطبيعي.

- (١ – ٥) حدود البحث

- تتمثل حدود البحث في اعتماد هذه الدراسة علي انتاج واستهلاك الغاز الطبيعي في مصر والبيانات متوفرة من ١٩٩٠ حتي ٢٠١٥ .

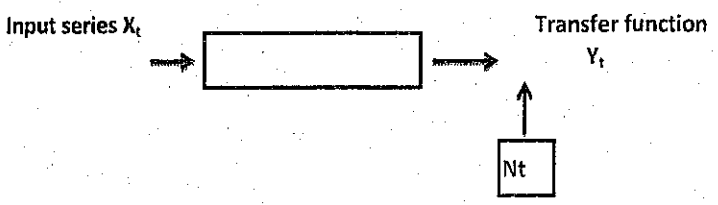
- الفصل الثاني : مكون من مبحثين الأول : نموذج الدالة التحويلية

(٢ – ١ – ١)مفهوم نموذج دالة التحويل Concept of Transfer Function Model :

نفترض أن لدينا سلسلتين زمنيتين X_t و Y_t ، السلسلة X_t تسمى سلسلة المدخل Input series والسلسلة Y_t تسمى سلسلة المخرج Output series ، ويفرض أن سلسلة المدخل تؤثر على سلسلة المخرج من خلال علاقة أو دالة تسمى الدالة التحويلية الثنائية ، وتتعرض Y_t بالإضافة إلى تأثير X_t لمتغيرات أخرى معروفة يضمن تأثيرها كلها فيما يسمى بمتغير الضجة ويرمز لها بالرمز N_t .

والشكل (2-1) يوضح نموذج مبسط للدالة التحويلية (إبراهيم ٢٠١٥)

والشكل (2-1) يوضح نموذج مبسط للدالة التحويلية :



وإذا كانت سلسلة المدخل وسلسلة المخرج غير مستقرة نقوم بتحويلها إلى سلاسل زمنية مستقرة بإجراء أى فروق مطلوبة وأى تحويلات لجعلها مستقرة فى المتوسط والتباين.

(٢ – ١ – ٢) هنالك صيغتان للدالة التحويلية :

♦ الصيغة الأولى :

هذه الصيغة مفيدة فى توضيح فكرة الدالة التحويلية الأساسية وتأخذ الشكل التالى :

$$Y_t = V_0X_t + v_1 X_{t-1} + \dots + V_kX_{t-k} + N_1 \rightarrow (1\text{-}2)$$

$$= (V_0 + V_1\beta + \dots + v_k \beta^k) X_t + N_t$$

$$= v(\beta)X_t + N_t$$

حيث :

- Y_t : سلسلة المخرج .
- X_t : سلسلة المدخل .
- N_t : الضجة .
- K : رتبة لدالة التحويلية .
- v_1, v_2 : أوزان الدالة التحويلية .

حيث تشير $v(\beta)$ إلى دالة التحويل الخاصة بنماذج بوكس – جنكيز حيث أن :

$$v(\beta) = \sum_{j=0}^{\infty} v_j \beta_j \rightarrow (2-2)$$

إن بوكس – جنكيز قاموا بتسمية النموذج (2-1) بنموذج دالة التحويل لأن السلاسل X_tو N_t تتبع بعض خصائص نماذج ARIMA

♦ الصيغة الثانية :

يفترض فى الصيغة الثانية أنه قد تم إجراء أى فروق مطلوبة وأى تحويلات مطلوبة على السلاسل N_t, Y_t, X_t لجعلها مستقرة من حيث المتوسط والتباين ، حيث يرمز للسلاسل بعد التعديل بـ n_t, y_t, x_t بالترتيب .

وهنالك بعض الصعوبات الموجودة فى نموذج دالة التحويل وهى أن المعلومات الموجودة بـ y_t, x_t هى محدودة ونهائية ولكن دالة التحويل $v(\beta)$ ربما تحتوى على عدد غير محدود من المعلمات وتم وضع الصيغة بحيث تتطلب عدداً أقل من المعلمات خاصة عندما تكون k كبيرة فى (2-1) .

وفق هذه الصيغة تكون الدالة التحويلية :

$$y_t = \frac{\omega(\beta)}{\delta(\beta)}\beta^b x_t + \frac{\theta(\beta)}{\phi(\beta)}a_t \rightarrow (3-2)$$

او

$$y_t = \frac{\omega(\beta)}{\delta(\beta)}\beta^b x_t + n_t \rightarrow (4-2)$$

◄ حيث :

$$\omega(\beta) = \omega 0 - \omega 1\beta - \ldots - \omega s\beta^s$$

$$\delta(\beta) = 1 - \delta_1\beta - \ldots - \delta_r\beta^r$$

$$\theta(\beta) = 1 - \theta_1\beta - \ldots - \theta_q\beta^q$$

$$\Phi(\beta) = 1 - \Phi 1\beta - \ldots - \Phi_p\beta^p$$

نلاحظ أن $\delta(\beta)$ و $\omega(\beta)$ يحلان محل $v(\beta)$ فى تحديد العلاقة بين السلسلتين الزمنيتين ، و $\theta(\beta)$ و $\Phi(\beta)$ مشغلاً المتوسط المتحرك والانحدار الذاتى المطلوبين لتخليص n_t من أثر هاتين العمليتين لتبقى فقط الضجة البيضاء a_t .

أما المعلمات q, p, s, r, b فتفسر كما يلى :

- (b) : تعنى أن التأخير delay (الزمن الميت) أو الفترة (عدد الوحدات الزمنية) قبل أن تبدأ x فى التأثير على y هو b وحدة زمنية ، وعلى هذا فإن سيكون تأثيرها الأول على y_{t+b} و x_{t-b} تؤثر أولاً على y_t وهكذا .

- (r) : تعنى أن y تتأثر بقيمتها السابقة حتى إبطاء r .

  أى y_t تتأثر بـ y_{t-1}, y_{t-2}, ..., y_{t-r} .
- (s) : تعنى أن القيمة الجديدة لـ x ستستمر فى التأثير على y لعدد s من الفترات الزمنية. أو بمعنى آخر y_t تتأثر بالقيم من x_{t-b} وحتى x_{t-b-s} .
- (q) : رتبة المتوسط المتحرك .
- (p) : رتبة الانحدار الذاتى .

وبصورة عامة قد تتأثر سلسلة المخرجات بسلاسل مدخلات متعددة فمن السهل توسعة النماذج أحادية الإدخال Single-Input series إلى نماذج متعددة الإدخال multi-input series ، بافتراض أن عدد متغيرات الإدخال فى النظام هو j فإن نموذج دالة التحويل ذات الإدخال المتعدد هو (البشير، ٢٠١٤) :

$$y_t = \frac{\omega_1(B)}{\delta_1(B)}x_{1,t-b_1} + \frac{\omega_2(B)}{\delta_2(B)}x_{2,t-b_2} + \cdots + \frac{\omega_j(B)}{\delta_j(B)}x_{j,t-b_j} + \frac{\theta(B)}{\phi(B)}a_t \quad (2-2)$$

(٢ – ١ – ٢) أوزان دالة التحويل (دالة الاستجابة النبضة) Impulse Response Function :

يمكن كتابة نموذج دالة التحويل (2-1) بالشكل المختزل التالى :

$$y_t = \nu(\beta)x_t$$

$$\nu(\beta) = \sum_{k=0}^{\infty} \nu_k \beta^k$$

$0, \nu_1, \nu_2, \ldots, \nu_k$ تعرف على أنها أوزان الدالة التحويلية (دالة الاستجابة النبضية) لسلسلة المدخل x_t وتمثل هذه الأوزان الأثر الذى يحدث على y_t نتيجة لتغير x_t بوحدة واحدة.

هذه الأوزان تزودنا بمقياس لكيفية تأثير سلسلة الإدخال فى سلسلة الإخراج ويرتبط الوزن بـ (التأخير الزمنى) بمعنى آخر أن V_0 مقياس لكيفية تأثير الاستجابة النبضية الحالية لسلسلة المخرجات بالقيمة الحالية لسلسلة المدخلات ، v_1 مقياس لكيفية تأثير الاستجابة النبضية الحالية لسلسلة المخرج بقيمة سلسلة المدخل لفترة زمنية واحدة ، v_2 مقياس لكيفية تأثير الاستجابة النبضية الحالية لسلسلة المخرج بقيمة سلسلة المدخل لفترتين زمنيتين ... إلخ

يسمى نموذج دالة التحويل بالنموذج الثابت إذا كانت سلسلة أوزان الاستجابة النبضية قابلة للجمع المطلق أى أن :

$$\sum_{k=0}^{\infty} |V_k| < \infty$$

ويسمى بالنموذج السببى (causal) إذا كانت $V_k = 0$ لجميع قيم $K < 0$ ، وعليه فإن النظام لا يستجيب لسلسلة المدخل حتى يتم تطبيقها على النظام فعلياً . (A. Yaffee, Robert, 2008)

(٢ – ١ – ٤) دالة الارتباط التقاطعى :

فى نموذج ARIMA المنفرد يعتبر معامل الارتباط الذاتى المفتاح الرئيسى لأنه يساعد فى تحديد شكل النموذج ، وفى نموذج دالة التحويل يؤدى الارتباط الذاتى دوراً ثانوياً ، بينما يؤدى الارتباط التقاطعى فى نموذج دالة التحويل دوراً رئيسياً حيث أنه يقيس الارتباط بين السلسلتين الزمنيتين سلسلة المدخل وسلسلة المخرج من جانب قوة العلاقة بين السلسلتين ومن جانب اتجاه العلاقة ، والارتباط التقاطعى لا يقيس قوة واتجاه العلاقة بين السلسلتين فحسب بل يوضح الصورة الكاملة للعلاقة بين سلسلة المدخل والمخرج خلال الفترات الزمنية المختلفة .

إذا كان لدينا سلسلة المخرج y_t وسلسلة المدخل x_t حيث أن $t = 0, \pm1, \pm2, \pm3, \ldots$ فإن التغير المقطعى (Cross covariance) بين x, y بإبطاء k ويرمز له بالرمز $C_{xy}(k)$ هو : (Douglas, 2008)

$$C_{xy}(k) = E\left[(x_t - m_s)(y_{t+k} - m_y) \right] \rightarrow (6\text{-}2)$$

$$k = 0, +1, +2, \ldots$$

$$m_x = E(x_t)$$

$$m_y = E(y_t)$$

فإن الارتباط التقاطعى هو :

$$P_{xy} = \frac{C_{xy}}{\sigma_x \ \sigma_y} \rightarrow (7\text{-}2)$$

◄ حيث :

- P_{xy} : الارتباط التقاطعى .
- C_{xy} : التغاير التقاطعى .

- $\sigma_x \sigma_y$: الانحرافات المعيارية للسلسلتين x و y .

(٢ – ١ – ٥) العلاقة بين نموذج دالة التحويل ودالة الارتباط الذاتي :

يمكن كتابة موذج دالة التحويل عند الزمن (t+k) بالصيغة التالية :

$$y_{t+k} = v_0 X_{t+k} + v_1 x_{t+k-1} + v_2 x_{t+k-2} + \dots + n_{t+k} \rightarrow (8\text{-}2)$$

بوضع $m_x = 0$, $m_y = 0$

ويضرب x_t فى طرفى النموذج (8-2) وبعد أخذ التوقع يصبح لدينا :

$$C_{xy}(k) = v_0 C_{xy}(k) + v_1 C_{xx}(k-1) + \dots \rightarrow (9\text{-}2)$$

بوضع $C_{xn}(k) = 0$ for all k يصبح لدينا :

$$P_{xy}(k) \frac{\sigma_x}{\sigma_y} [\, v_0 p_x(k) + v_1 p_x(k-1) + v_2 p_x(k-2) + \dots \,] \rightarrow (10\text{-}2)$$

إذا كانت سلسلة المدخل هى ضوضاء بيضاء إذن :

$$p_x(k) = 0 \text{ for } k \neq 0$$

عليه فإن أوزان الدالة التحويلية v_k تعطى بالصيغة التالية :

$$v_k = \rightarrow \frac{P_{xy}(k)\sigma_x}{} \quad (11\text{-}2)$$

◄ حيث :

- v_k : أوزان الدالة التجريبية .
- $P_{xy}(k)$: الارتباط التقاطعى بين السلسلتين x, y .
- σ_x ، σ_y : الانحرافات المعيارية للسلسلتين x, y .

إذن فإن دالة الاستجابة النبضية تتناسب مع دالة الارتباط التقاطعى فى نموذج دالة التحويل .(A. Yaffee, Robert, 2008)

(٢ – ١ – ٦) مراحل بناء نموذج دالة التحويل :

تمر عملية بناء نموذج الدالة التحويلية بنفس مراحل بناء نموذج أريما ، وهى تحديد النموذج وتقدير المعلمات وإجراء اختبار تشخيصى للنموذج مع الفارق فى أن المرحلة تمر أيضاً بعملية تنقية مكثفة للسلسلتين الزمنيتين من المؤثرات المعروفة .

فإذا كانت كل من سلسلة المدخل x_t وسلسلة المخرج y_t بشكلها الخام فإن خطوات بناء نموذج الدالة التحويلية يمكن تلخيصها فى الخطوات التالية :

فإذا كانت كل من سلسلة المدخل x_t وسلسلة المخرج y_t بشكلها الخام فإن خطوات بناء نموذج الدالة التحويلية يمكن تلخيصها فى الخطوات التالية :

*** تجهيز سلسلة المدخل xt وسلسلة المخرج yt :**

ويعنى ذلك إجراء الفروق اللازمة لتحقيق الاستقرار فى المتوسط ، وإجراء التحويلات اللازمة لتحقيق الاستقرار فى التباين . كذلك تتم فى هذه الخطوة إزالة أى تأثير موسمى من السلسلتين إن وجد. (برى، ٢٠٠٢)

*** إجراء تبيض مسبق Prewhitening لكل من سلسلة المدخل وسلسلة المخرج :**

تبيض السلسلة x_t يقصد به بناء نموذج أريما يمثلها وتطبيقه على x_t للحصول سلسلة البواقى α_t وذلك كالآتى :

• نفترض أن لدينا سلسلة المدخل x_t ونموذج ARIMA لها كالتالى :

$$\omega(\beta) = \omega 0 - \omega 1\beta - \dots - \omega s\beta^{s}$$

$$\delta(\beta) = 1 - \delta_1\beta - \dots - \delta_r\beta^{r}$$

$$\theta(\beta) = 1 - \theta_1\beta - \dots - \theta_q\beta^{q}$$

$$\Phi(\beta) = 1 - \Phi 1\beta - \dots - \Phi_p\beta^{p}$$

$$\Phi_x(\beta)x_t = \Phi_x(\beta)\alpha_t \rightarrow (12 - 2)$$

حيث أن α_t ضجة بيضاء إذن :

$$\alpha_t = \frac{\Phi_x(\beta)}{x_t} \rightarrow (13 - 2)$$

تسمى سلسلة المدخل α_t بالسلسلة المبيضة .

كذلك سوف يتم تبيض سلسلة المخرج y_t بنفس الطريقة وكما يلى :

حيث أن α_t ضجة بيضاء إذن :

$$\frac{\beta_t(\beta)}{x t \theta_x(\beta)} \quad \rightarrow (14-2)$$

الهدف من ذلك التبيض هو تنقية السلاسل الزمنية x_t, y_t بإزالة أى نمط معروف ناتج عن عملية انحدار ذاتى أو متوسط متحرك ، فلا تبقى فيها سوى ضجة بيضاء هى α_t ، β_t .

وستكون العلاقة بين α_t ، β_t خالية من تأثيرات عمليات الانحدار الذاتى والمتوسط المتحرك.

(٢ – ١ – ٧) حساب الارتباطات التقاطعية :

فى هذه الخطوة نحسب الارتباطات التقاطعية بين سلسلة المدخل α_t وسلسلة المخرج β_t بإبطاءات مختلفة بين α_t ، β_t .

إن الارتباط التقاطعى بإبطاء k بين α_t ، β_t يعطى بالصيغة التالية :

حيث أن α_t ضجة بيضاء إذن :

$$= \rho\alpha_t\beta_t \xrightarrow[\sigma_\alpha \sigma_\beta]{C\alpha_t\beta_t} \quad (15-2)$$

حيث :

- $\rho\alpha_t\beta_t$: الارتباط التقاطعى بين السلسلتين α_t ، β_t .
- $C\alpha_t\beta_t$: التغاير التقاطعى بين α_t ، β_t .
- $\sigma\alpha_t\, \sigma\beta_t$: الانحرافات المعيارية للسلسلتين α_t ، β_t .

ويعطى فى الواقع الارتباط بين قيم α_t فى الزمن t وقيم β_t التى تبعد عنها زمنياً بـ k وحدة من الزمن t + k.

إذا كان السلسلتان ضجة بيضاء فإن الارتباط المقطعى سيكون متوسطه صفر وتباينه $\frac{1}{n}$ ،أما إذا كانت إحداهما فقط ضجة بيضاء ، فإن الخطأ المعيارى (١٩٤٦) Barilet للارتباط المقطعى بإبطاء k يكون تقريباً $\sqrt{\frac{1}{n-k}}$.

(٢ – ١ – ٨) تقدير مباشر لأوزان الدالة التحويلية :

يقصد بالأوزان هنا $v_1, v_2, \ldots, v_n$ فى النموذج (2-1) .

يمكن كتابة النموذج (2-1) بدلالة السلاسل x, y, n التي تأجريت عليها الفروق لجعلها مستقرة (بافتراض b = 0) كالتالى :

$$y_t = v(\beta)x_t + n_t \rightarrow (16 - 2)$$

إذا قمنا بتبييض السلال الثلاث باستخدام التحويلة $\frac{\emptyset_x(\beta)}{\theta..(\beta)}$ أى وضعنا :

$$\frac{\emptyset_x(\beta)}{\theta..(\beta)} y_t = v(\beta) \frac{\emptyset_x(\beta)}{\theta..(\beta)} x_t + \frac{\emptyset_x(\beta)}{\theta..(\beta)} n_t$$

نتحصل على :

$$\beta_t = v(\beta)\alpha_t + \alpha'_t$$

بضرب الطرفين فى α_{t-k} وأخذ التوقع يصبح لدينا كالتالى :

$$E(\alpha_{t-k}\beta_t) = v_0 E(\alpha_{t-k}\alpha_t) + v_1 E(\alpha_{t-k}\alpha_{t-1}) + \dots + v_k E(\alpha_{t-k}\alpha_{t-k}) + E(\alpha_{t-k}\alpha_t)$$

ويما أن الضجة a يفترض أنها مستقلة عن α ، ويما أن $\alpha's$ مستقلة عن بعضها البعض ، فإن جميع الحدود فى الطرف الأيمن ستكون أصفاراً ما عدا الحد قبل الأخير حيث يساوى تباين α مضروباً فى v_k . أما الطرف الأيسر فهو التغاير المقطعى وبالتالى :

$$vk = \frac{C\alpha\beta(K)}{\sigma_u^2} \quad \frac{C\alpha\beta(K)\,\sigma\beta}{\sigma\beta\sigma\alpha\sigma\alpha}$$

$$vk = \frac{C\alpha\beta(K)}{\sigma_u^2} \quad \frac{C\alpha\beta(K)\,\sigma\beta}{\sigma\beta\sigma\alpha\sigma\alpha}$$

$$vk = \rightarrow (18 - 2) \quad \frac{\rho\alpha\beta(K)\,\sigma\beta}{\sigma\beta}$$

حيث :

- v_k : أوزان الدالة التحويلية .
- $\rho\alpha\beta(K)$: الارتباط التقاطعى بين السلسلتين α_t ، β_t .
- $\alpha\beta$: الانحراف المعيارى للسلسلة αt .
- σu : الانحراف المعيارى للسلسلة β_t .

وبالتالى يمكن تقدير الوزن ذو الرتبة k بضرب مقدر الارتباط المقطعى بين αt و β_t فى الانحراف المعيارى للسلسلة β_t والقسمة على الانحراف المعيارى للسلسلة αt(A. Yafee, Robert, 2008)

(٢-١-٩) تحديد القيم r, s, b لنموذج الدالة التحويلية :

إن تحديد القيم r, s, b ليس سهلاً ، ولكن يمكن الاستهداء ببعض القواعد عند تحديد هذه القيم وهى:

- لتحديد القيمة b (الزمن الميت) ننظر إلى قيم أوزان الدالة التحويلية التى تحسب من المعادلة (2-17) فى الخطوة السابقة ، حيث أن b تساوى عدد الأوزان التى تساوى صفراً بعد الوزن v_0 (الموسوى، لفتة، ٢٠١٢).
- نفحص الارتباطات الذاتية المقطعية التى تحسب من المعادلة (2-15) :

١- فإذا كانت الارتباطات المقطعية غير معنوية حتى الإبطاء m حيث أصبحت معنوية نأخذ b = m .

٢- إذا لم يكن هناك نمطاً معيناً للارتباطات المقطعية بعد الإبطاء m وحتى الإبطاء m + a نضع s = a .

٣- إذا ظهر نمط محدد بعد m + a وحتى m + a + c نضع r = c . (البشير، ٢٠١٤).

ثانيا : المبحث الثاني : الشبكات العصبية

(٢-١) الشبكات العصبية الاصطناعية Artificial Neural Networks :

سميت الشبكات العصبية بهذا الاسم كونها شبكة من وحدات الاتصالات الداخلية ، حيث إن هذه الوحدات مأخوذة من دراسة أنظمة الأعصاب الحيوية ، أو بعبارة أخرى إنها محاولة بناء ميكانيكية باستخدام الهيكليات التى تسلك نفس سلوك العصبونات البيولوجية .

ومن الجدير بالذكر أن موضوع الشبكات العصبية هو جزء من موضوع الذكاء الاصطناعى Artificial Intelligence (AI) ، الذى يشمل فروع كثيرة منها علم الفلسفة والرياضيات وعلم النفس التأملى Cybernetics وعلم الحاسوب ، حيث أن بعض مفردات الذكاء الاصطناعى تتعلق بفهم كيفية عمل الدماغ أو العقل البشرى ومحاولة محاكاة نموذجه باستخدام الحاسوب ، أو بعبارة أخرى نمذجة الذكاء كيان مفرد Single entity عن طريق محاولة بناء نموذج حاسوبى له نفس السلوك فيما يتعلق بهيكلية الدماغ حيث أنه يتألف من ملايين

عصبونات الاتصالات الداخلية لذلك كان استخدام الشبكة العصبية الاصطناعية محاولة إنتاج نشاط مشابه لعمل الدماغية الاصطناعية .

(٢-٢-١) خصائص الشبكات العصبية الاصطناعية : (التعيبى، ٢٠٠٢)

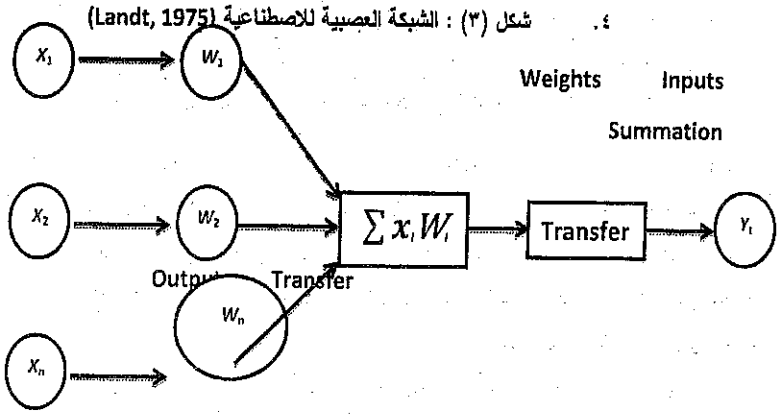
تتميز الشبكات العصبية بالعديد من الخصائص أهمها :

١. تعتمد على أساس رياضى قوى ز
٢. تمثل إحدى تطبيقات تكنولوجيا التشغيل الذكى للمعلومات التى تقوم على محاكاة العقل البشرى .
٣. تقبل أى نوع من البيانات الكمية أو النوعية .
٤. لها القدرة على تخزين المعرفة المكتسبة من خلال الحالات التى يتم تشغيلها على الشبكة .
٥. يمكن تطبيقها فى العديد من المجالات العلمية المختلفة .

(٢-٢-٢) مكونات ومنظومة الشبكات العصبية الاصطناعية : (Landt, 1997)

تتكون الشبكة العصبية من مجموعة مركبة من عدة عناصر معالجة يطلق عليها اسم نيرون (Neuron) وهذه العناصر لها القدرة على إجراء العمليات الحسابية من خلال عدة خطوات :

١. معالجة البيانات تمر أولاً من خلال عناصر المعالجة البسيطة (Neuron) .
٢. تمر الإشارات بين الأعصاب عبر خطوط الربط الناقلة .
٣. ومن خطوط الربط يتم إعطاء وزن معين لكل مدخل معين ويضرب مع الإشارة الداخلة إلى (Neuron) . ويوضح شكل (٣) نموذجاً بسيطاً للشبكة العصبية الاصطناعية .
٤. شكل (٣) : الشبكة العصبية للاصطناعية (Landt, 1975)



ويتضح من الشكل السابق أن الشبكة العصبية تتكون من ثلاثة شرائح على النحو التالى :

١. شريحة المدخلات (Input Layer).

٢. شريحة مخفية (Hidden Layer).

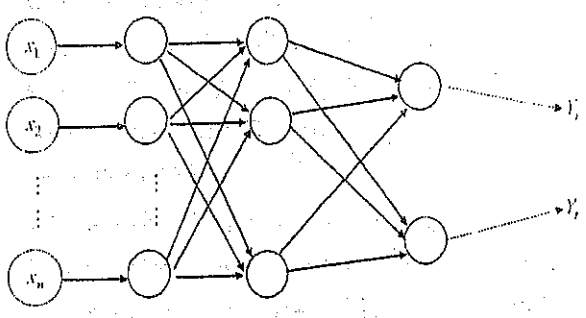
٣. شريحة المخرجات (Output Layer).

(٢-٢-٢) أنواع الشبكات العصبية الاصطناعية : (الكفراوى Mtchell 1997 & 2002)

تنظم الشبكة العصبية الاصطناعية لعدد هيئات مختلفة ، بمعنى توصيل الخلايا العصبية بعدة طرق مختلفة منها :

(١) الشبكات العصبية ذات التغذية الأمامية : (Feed Forward Neural Networks)

وهى الشبكات التى يخلو تركيبها من وجود حلقة مغلقة من الترابطات بين الوحدات المكونة لها . وتعد هذه الشبكات من أكثر الشبكات استخداماً؛ حيث تتكون الشبكة من هذا النوع من طبقتين على الأقل ، كما تتواجد فى كثير من الأحيان طبقات مخفية (Hidden Layer) بين طبقة المدخلات وطبقة المخرجات، وتنتقل العمليات الحسابية فى اتجاه واحد إلى الأمام من طبقة المدخلات إلى طبقة المخرجات عبر الطبقات المخفية كما يتضح من شكل (4) التالى :



شكل (4) : نموذج لشبكة عصبية ذات تغذية أمامية، (Landt, 1975)

(٢) الشبكات العصبية ذات التغذية المرتجعة : (Feed Back Neural Networks)

وهى الشبكات التى تجد لمخرجاتها طريقاً خلفياً مرة أخرى لتصبح مدخلات لكى تعطى أفضل النتائج الممكنة .

(٣) الشبكات العصبية ذات الترابط الذاتى : (Auto Associative Neural networks)

وهى الشبكات التى تلعب كافة العناصر المكونة لها دوراً نموذجيا ، يتمثل فى استقبال المدخلات وبث المخرجات فى نفس الوقت .

(٢-٢-٤) التصميم الهيكلى لبناء الشبكة العصبية الاصطناعية :

التصميم الهيكلى لبناء الشبكات العصبية الاصطناعية (ANN) لنظم المعلومات التقليدية ، أى المعروفة وفقاً للحاسبات والبرامج فى حقل الشبكات العصبية ، ويتضمن ذلك التصميم الخطوات التالية :

- الخطوة الأولى : يتم تجميع البيانات التى تستخدم فى تدريب أو اختبار الشبكة .
- الخطوة الثانية : يتم تعريف البيانات الخاصة بالتدريب للشبكة ، ووضع خطة التدريب والتعلم .
- الخطوة الثالثة : بناء هيكل الشبكة وتحديد تركيب الشبكة فى عدد المدخلات والطبقات ونوع الشبكة .
- الخطوة الرابعة : اختيار طريقة التعلم حسب ما هو متاح من أدوات أو تطوير الشبكة أو القرارات .
- الخطوة الخامسة : وضع قيم للأوزان والمتغيرات ، ثم بعد ذلك يتم تعديل القيم فى الأوزان عن طريق التغذية المرتدة والعكسية .
- الخطوة السادسة :" تحويل البيانات إلى النوع المناسب للشبكة ويتم ذلك عن طريق كتابة معادلة لتجهيز البيانات والبرامج الجاهزة .
- الخطوة السابعة والثامنة :
- وفيهما تتم عمليتى التدريب والاختبار من خلال تكرار عرض المدخلات والمخرجات المرغوبة إلى الشبكة ، ومنها مقارنة القيم الفعلية مع القيم المحسوبة ، ثم حساب الفرق (الخطأ) ، ثم تعديل الأوزان لتقليل الفرق حتى يصبح فرقاً مقبولاً .
- الخطوة الأ×يرى : وهنا تستطيع الشبكة الوصول إلى النتائج المرغوبة من خلال استخدام مدخلات التدريب ، وبهذا يمكن الاعتماد على الشبكة فى الاستخدام كنظام مستقل قائم بذاته أو كجزء من النظام .

(٢-٢-٥) مزايا وعيوب الشبكات العصبية :

يتميز أسلوب الشبكات العصبية عن غيره من البدائل الإحصائية فى عملية تحليل البيانات بالمزايا التالية :

(٢) مرحلة الانتشار الخلفى للخطأ :

بعد انتقال القيم إلى طبقة المخرجات التى تم حساب قيمتها من خلال الخطوات السابقة ، يتم إجراء مقارنة بين القيم المحسوبة والقيم المرغوبة (حساب الخطأ) من خلال الفرق بين قيم تلك المخرجات وذلك من خلال معادلة الخطأ التالية :

$$E = (xj - yj)$$

حيث :

- xj← المخرجات المرغوبة من الشبكة .
- yj← المخرجات المحسوبة من الشبكة .

وبعد ذلك يتم تصحيح الوزن وتعديله من خلال عملية التعلم التى تتم على الشبكة ، وذلك من خلال المعادلة التالية (Haykins, 1999) :

$$Wi(Final) = Wi + \alpha.\beta.xi$$

وتعبر α عن معدل التعلم والذى يتم وضعه عند أدنى مستوى تعلم ، بينما تعبر β عن الفرق بين القيمة المحسوبة والقيمة المرغوبة ($\beta = x_i - y_i$) ، وذلك لحساب الخطأ وتسمى بالمرحلة التراجعية (Backward) وتكرر هذه الخطوات فى الشبكة لعدة مرات بخطوة أمامية وخطوة تراجعية (مرحلة التغذية الأمامية – مرحلة الانتشار الخلفى) تسمى بدورة (Epoch) .

◘ البعد المطمور (Finding the Embedding Dimension) :

يعتمد البعد المطمور على مفهوم (Cross Validation) حيث يتم تقسيم بيانات التدريب إلى عينتين منفصلتين ، عينة للتحقيق وعينة جديدة للتدريب ، ثم يتم مقارنة الخطأ وبعد ذلك يتم اختيار أفضل نموذج بعد المدخلات لذلك النموذج ، وعندئذ يكون هذا تخمين جيد للبعد المطمور .

(٢ – ٢ – ٧) العوامل المؤثرة على كفاءة الشبكة العصبية الاصطناعية :

Artificial neural network performance factor

هناك عوامل تؤثر على كفاءة الشبكة العصبية الاصطناعية ، حيث أن جودة التنبؤات المستقلية لظاهرة معينة يمكن الحصول عليها من الشبكة العصبية الاصطناعية بحيث تعتمد وبشكل أساسى على مدى كفاءة تدريب الشبكة العصبية الاصطناعية وذلك للبيانات التاريخية لتلك الظاهرة ، وهذه العوامل هى :

١. بمجرد أن تقوم الشبكة بتلقى البيانات فإن أدائها لا يتأثر كثيراً عند إدخال أى مجموعة إضافية من البيانات الجديدة والتى لم تكن موجودة فى البداية .

٢. لا تتطلب وجود افتراضات قوية بين بشأن نوعية البيانات والعلاقة بين المتغيرات كما فى الأساليب الإحصائية الأخرى .

٣. لا تتطلب توافر عدد كبير من البيانات التجريبية .

٤. ليس هناك حاجة لترتيب المتغيرات حسب أهميتها ، وذلك حيث أن الشبكة تضع تلقائياً أوزان خاصة بها حسب طبيعة البيانات المدخلة إليها .

٥. تعطى قرارات واضحة فى النماذج المعروفة ، كما يمكنها معالجة مشكلات اتخاذ القرارات والتى تتصف بمدخلات غير واضحة (Stochastic) .

ويعاب عليها أنها لا تتعامل مع المشاكل ذات المتغيرات التى لا يمكن صياغتها كمياً ، كما أنها تعطى نتائج فى بعض الحالات قد لا نستطيع تقديم تفسيراً منطقياً لها (علاقات هوائية).

(٢-٢-٦) استخدام الشبكات العصبية فى التنبؤ :

بفرض وجود سلسلة زمنية تحتوى على المشاهدات $X_n, X_{n-1}, ..., x_2, x_1$ ، وللتنبؤ بالقيم المستقبلية للمتغير بمعنى إيجاد القيم X_{n+1}, X_{n+2} ، وكأن السلسلة محددة بمعنى وجود عدد حقيقى P الذى يدعى البعد الممطور (The Embedding Dimension) والدالة f بحيث تكون $t > P$.

$$X(t) = f(X_{t-1}, X_{t-2}, ..., X_{t-p})$$

وبالتالى يتكون التنبؤ من N قيمة لسلسلة زمنية معطاة ، ولإيجاد قيم p, f وبفرض أن p معروفة ونرغب فى إيجاد f ومن المعروف أن الشبكات العصبية الاصطناعية مكونة من عدد كبير من عناصر بسيطة غير خطية مترابطة فإن هناك ثلاثة أنواع من الوحدات :

١. وحدات إدخال : والتى تهيئ للقيم السابقة فى السلسلة الزمنية $X_{t-1}, X_{t-2}, ..., X_{t-p}$ حيث p تمثل البعد المطمور .

٢. وحدات مخفية : والت تستخدم لحفظ التمثيل الداخلى للسلسلة الزمنية.

٣. وحدات إخراج : والتى تعطى مخرجات الشبكة العصبية فى حالة وحدة إخراج فقط لإرجاع $X(t)$.

ومن ثم حساب الدالة اللوجيستية $Y_t = \frac{1}{1 + e^{-Y}}$ التى تستخدم لتحويل البيانات إلى الخطية

الأول : عدد المتجهات أو النماذج الداخلة إلى الشبكة :

Number of Exemplars in Neural network

يؤثر عدد المتجهات أو النماذج التى يتم إدخالها إلى الشبكة العصبية عند إجراء عملية التدريب أو التعليم على أداء الشبكة حيث أنه يرتبط وبشكل مباشر بعدد المتغيرات المستقلة (الجزء الأول من المتجه) والتى تمثل المدخلات لعقد الشبكة، فعندما يكون عدد المتجهات أو النماذج الداخلة إلى الشبكة أكبر من عدد المدخلات أو المتغيرات المستقلة لكل نموذج تقوم الشبكة العصبية باستخلاص نموذج من تلك البيانات لتمثيل كل العلاقات والخصائص ذات الطبيعة المعقدة التى استطاعت الشبكة تعلمها من تلك البيانات ، وبمعنى آخر إذا كانت مدخلات الشبكة العصبية أى النماذج المقدمة لها على درجة من التعقيد فيجب إدخال عدد من تلك النماذج أكبر من عدد المتغيرات المستقلة لكل نموذج عند تدريب الشبكة على تلك المدخلات .

الثانى : عدد العقد المخفية **Number of Hidden Nodes**

يتم تحديد العدد المناسب من العقد المخفية للشبكة عند تدريبها من قبل المستخدم ن وللبدء بتدريب الشبكة لابد من إعطاء تقدير أولى لعدد العقد المخفية حيث أن الاستراتيجية المثلى لتحديد عدد العقد الخفية تتمثل باختيار عدد قليل من العقد المخفية عند البدء بتدريب الشبكة العصبية ، أى أن يبدأ تدريب الشبكة بعقدتين مخفيتين مثلاً ثم نقوم بمراقبة خطأ التدريب الذى يتناقص مع استمرار تدريب الشبكة ، وعندما يصل إلى حد معين يتوقف فيه عن التناقص ، نقوم بزيادة عدد العقد المخفية بعقدة واحدة ليبدأ تدريب الشبكة من جديد بعدد جديد من العقد المخفية م عالمراقبة المستمرة للخطأ .

إن هذه العملية المتمثلة بإضافة عقدة مخفية واحدة فى كل مرة والبدء بتدريب الشبكة من جديد تكرر لحين الوصول إلى القيمة المسموح بها للخطأ ، وعند الوصول إلى تلك القيمة ينتهى تدريب الشبكة ليمثل التدريب الأمثل ويكون عدد العقد المخفية فى الشبكة هو العدد الأمثل.

الثالث : عدد الطبقات المخفية **Number of Hidden layers**

لا يقل هذا العامل أهمية عن العوامل السابقة ، وكما هو الحال فى العامل السابق (عدد العقد المخفية) يفضل البدء بعدد قليل من الطبقات المخفية ، مبدئياً يتم البدء بتدريب الشبكة العصبية بطبقة مخفية واحدة وعدد قليل من العقد المخفية حيث تتم زيادة عدد العقد المخفية بالتدريج لحين الوصول إلى أقل خطأ ممكن ، ولكن الذى يحدث فى بعض الحالات وعلى الرغم

من زيادة عدد العقد المخفية فى الطبقة الواحدة إن تدريب الشبكة العصبية لا يتم بالشكل الأمثل ، أى لا يمكن الوصول إلى أقل خطأ ممكن فى تدريب الشبكة العصبية ، ويرجع ذلك إلى أن بعض العقد المخفية الموجودة فى الطبقة الواحدة تستطيع تعلم خصائص مجموعة معينة من البيانات التى تم إدخالها إلى الشبكة بشكل جيد ، فى حين تفقد بقية العقد المخفية القدرة على الاستجابة والتعلم لخصائص البيانات ، بمعنى آخر ، من الصعب تكوين شبكة عصبية اصطناعية بمجموعة من العقد المخفية ، بحيث أن كل عقدة مخفية تستجيب لتعلم خصائص مجموعة معينة من البيانات بشكل صحيح ، وللتغلب على هذه المشكلة يتم إضافة طبقة مخفية ثانية حيث تستطيع العقد المخفية الموجودة فى هذه المنطقة المضافة من استكمال عملية التعلم التى لتم تتم من قبل العقد المخفية فى الطبقة الأولى ن فإضافة هذه الطبقة الجديدة يصبح بمقدور الشبكة العصبية تعلم كل الخصائص والمزايا للبيانات التى تم إدخالها إلى الشبكة لغرض التعلم ، وكل ذلك بالطبع يعتمد على طبيعة البيانات التى تم إدخالها .

(٢ - ٢ - ٨) وحدات المعالجة (العصبونات) :

وحدات المعالجة أو العصبونات هى الوحدات التى تقوم بعملية معالجة المعلومات فى الشبكة العصبية وهى تشكل المكونات الأاسية التى تتألف منها كل طبقات الشبكة العصبية . وتتصل هذه الوحدات بطرق مختلفة بواسطة الواصلات البينية لتعطى الشكل العام أو البنية المعمارية للشبكة العصبية . (يعقوب، ٢٠٠٩)

تتبع عناصر المعالجة نظام المعالجة المتوازية فى إجراء الحسابات المسندة إليها أو معالجة البيانات وهى فى ذلك تتبع عمل العقل البشرى .

وتتألف أى وحدة معالجة أو عصبون من المكونات الأساسية التالية :

١) معاملات الأوزان . ٢) دالة الجمع .

٣) دالة التحويل . ٤) دالة الإخراج .

(١) معاملات الأوزان :

يعتبر الوزن هو العنصر الرئيسى للشبكات العصبية الاصطناعية فهى تمثل الروابط المختلفة التى يتم عبرها نقل البيانات من طبقة إلى طبقة أخرى . ويعبر الوزن عن القوة النسبية أو الأهمية النسبية لكل مدخل إلى عنصر المعالجة .

وتتعلم الشبكة من خلال ضبط الأوزان ويرمز للوزن بين عنصرى معالجة i و j بالرمز w_{ij} .

(٢) دالة الجمع :

إن أول عملية تقوم بها وحدة المعالجة هى حساب مجموع المدخلات الموزونة القادمة إلى الوحدة باستخدام دالة الجمع . حيث تقوم هذه الدال بحساب متوسط الأوزان لكل مدخلات وحدة المعالجة ، ويتم ذلك بضرب كل قيمة مدخلة فى وزنها المصاحب ومن ثم إيجاد المجموع لكل حواصل الضرب ، يعطى ذلك رياضياً كما يلى :

$$S_j = \sum_{i=1}^{n} x_i w_{ij} \rightarrow (33\text{-}2)$$

حيث :

- S_j : ناتج عملية الجمع لكل وحدة معالجة j .
- X_i : القيمة المدخلة القادمة من الوحدة i والداخلة إلى الوحدة j .
- w_{ij} : الوزن الذى يربط وحدة المعالجة j بالوحدة i الموجودة فى الطبقة السابقة .

(٣) دوال التحويل :

إن العملية الثانية فى وحدة المعالجة بعد عمل دالة الجمع هو تحويل ناتج الجمع إلى أحد القيم التى يفترض أن تكون ضمن نواتج الشبكة المرغوب بها .

وتتم هذه الخطوة باستخدام دالة تسمى بدالة التحويل حيث تقوم هذه الدالة بتحويل عملية الجمع الموزون فى الخطوة الأولى إلى قيمة محصورة فى مدى معين ويتم ذلك بمقارنة نتيجة الجمع مع قيمة معينة تسمى قيمة العتبة ويرمز لها بالرمز θ ليتحدد الناتج ومن أهمها:

(٢- ٢ -٩) توابع التحويل :

a- دالة سيقمويد Sigmoid Function .

b- دالة الخطوة Step Function .

c- الدالة الخطية Linear Function .

d- دالة الإشارة Sign Function

(a)دالة سيقمويد Sigmoid Function :

هذه الدالة تجعل المخرجات أو تحولها إلى قيمة محصورة بين (0 و 1) وتسمى فى هذه الحالة بدالة تنشيط سيقمويد الثنائى ، أو تحويل المخرجات إلى قيم بين (-1, +1) وتسمى بدالة سيقمويد ثنائى القطبية .

وتعتبر هذه الدالة من أكثر الدوال استخداماً خاصة فى خوارزميات تدريب شبكات الانتشار الخلفى . تأخذ هذه الدالة الصيغ الرياضية التالية :

$$f(x) = \frac{1}{1+e-s} \rightarrow (34\text{-}2)$$

(b) دالة الخطوة Step Function :

تعرف هذه الدالة بأنها من دوال الحد الفاصل وهذا التابع يجعل القيمة المخرجة فى وحدة المعالجة محصورة بين (0, 1) كالنظام الثنائى . وتأخذ هذه الدالة الصيغة الرياضية التالية :

$$f(x) = \rightarrow (35\text{-}2) \quad \begin{cases} s \geq \theta & \text{if} \quad 1 \\ s < \theta & \text{if} \quad 0 \end{cases}$$

(c) الدالة الخطية Linear Function :

توفر دالة التنشيط الخطى مخرجات تساوى المدخلات الموزونة لوحدة المعالجة وعادة تستخدم وحدات المعالجة بالدالة الخطية فى التقريب الخطى . تعطى هذه الدالة بالصيغة الرياضية التالية :

$$f(x) = S \rightarrow (36 - 2)$$

(d) دالة الإشارة Sign Function :

تعتبر هذه الدالة من دوال الحد الصلب وتستخدم عادة فى وحدات المعالجة لشبكات عصبية تستخدم فى التصنيف وتمييز الأنماط . وتستخدم هذه الدالة قيمة معينة تسمى θ .

تأخذ هذه الدالة الصيغة الرياضية التالية :

$$f(x) = \rightarrow (37\text{-}2) \quad \begin{cases} s \geq \theta & \text{if} \quad +1 \\ s < \theta & \text{if} \quad -1 \end{cases}$$

(٤) دالة المخرجات :

بعد أن تقوم دالة الجمع بعملية الجمع الموزون للمدخلات ومن ثم دالة التحويل بتحويل ناتج الجمع إلى ناتج قيمة محصورة فى مدى معين . فقد تكون المخرجات فى أغلب الأحيان مساوياً لناتج دالة التحويل . ولكن هناك بعض الشبكات تقوم وحدة المعالجة فيها بتعديل نتيجة دالة التحويل .

ثالثا المبحث الثالث : السلاسل الزمنية

(٢-٣) : السلاسل الزمنية : نماذج بوكس- جينكز

ظهر أسلوب تحليل السلاسل الزمنية بشكل واسع على يد العالمين G.Box, G. Jenkis فى كتابهما Times Series Analysis forecasting and control عام (1970) وأصبح هذا الأسلوب متاحاً فى حزم البرامج الجاهزة ونقلاً عن Chui Chang(1992) فإنه قد بدأت دراسة المقارنة العلمية لطرق السلاسل الزمنية مع البحث الذى قدمه كلاً من Granger and New bold(1974) وأول منافسة للمنهجيات كانت لكل من McGollister and Grange(1978) وكانت أكثر المقارنات شمولاً هى التى أجراها Makridaris et al., (1982 ويطلق عليها "منافسة أم" (M-Competition) حيث طبق سبعة خبراء أربع وعشرون طريقة تنبؤية للسلاسل الزمنية أما على (1001) سلسلة بيانات أو على عينة فرعية من (111) سلسلة بيانات وشملت السلسلة بيانات سنوية وربع سنوية وشهرية غطت جميع أنواع البيانات لعدة شركات وتوصل الخبراء إلى أنه لا توجد طريقة تنبؤ بمفردها أفضل من الطريقة البسيطة ، ولقد أعطى هذا الأسلوب تنبؤات ذات دقة عالية فى كثير من الدراسات .

(٢-٣-١) نماذج الانحدار الذاتى والمتوسطات المتحركة وخصائصها ARMA(p,q)

Autoregressive Moving average (p,q)

تعتمد النماذج الحديثة للسلاسل الزمنية على المبدأ الفلسفى القائل بأن الحاضر هو نتاج الماضى وهذا يعنى أنه يمكن التعبير عن المشاهدة الحالية Y_t كدالة خطية فى المشاهدات السابقة $Y_{t-p} \dots Y_{t-2}, Y_{t-1}$ بالإضافة إلى الأخطاء العشوائية الماضية $\varepsilon_{t-q} \dots \varepsilon_{t-2}, \varepsilon_{t-1}$ ومن أشهر هذه النماذج ما يعرف باسم نماذج الانحدار الذاتى والمتوسطات المتحركة ARMA(p,q) حيث تمثل :

- P رتبة الجزء الخاص بالانحدار الذاتى .
- Q رتبة الجزء الخاص بالمتوسطات المتحركة .

افترض بوكس وجنكنز أن نماذج الانحدار الذاتى والمتوسطات المتحركة ARMA مناسبة لوصف سلوك السلسلة الزمنية والتنبؤ بها ، وقدما طريقتهما فى نمذجة بيانات السلسلة الزمنية ، وتشترط هذه الطريقة سكون السلسلة الزمنية فى المتوسط (السلسلة لها اتجاه عام) بالإضافة إلى ثبات التباين (أى تكون التذبذبات حول المتوسط الثابت متسقرة طول السلسلة الزمنية) .

وعلى ذلك فإن أولى خطوات التحليل الحديث للسلسلة الزمنية هى التحقق من توفر هذين الشرطين وذلك عن طريق رسم المنحنى الزمنى للسلسلة الزمنية ، فإذا لاحظنا منه عدم تحقق أى من الشرطين فإنه يتم إجراء عدد من التحويلات على السلسلة الزمنية كالآتى :

إذا كانت السلسلة الزمنية غير ساكنة فى المتوسط فإنا نقوم بأخذ الفروق المتتالية لتسكينها ، ففى حالة كون معادلة الاتجاه العام من الرتبة الأولى فإننا نحصل على السلسلة الساكنة من خلال أخذ الفرق الأول وفى حالة كونها من الرتبة الثانية فإننا نأخذ الفرق الثانى للحصول على السلسلة الساكنة ونرمز عادة لعدد الفروق التى يتم أخذها لتسكين السلسلة بالرمز (d) .

إذا كانت السلسلة غير ساكنة فى التباين نقوم بإجراء تحويله مناسبة مثل الجذر التربيعى أو اللوغاريتم ، وإذا كانت السلسلة غير ساكنة فى كل من التباين والمتوسط نقوم بإجراء تحويله لتسكين التباين أولاً ثم نختبر هل السلسلة ساكنة فى المتوسط ، فإذا كانت غير ساكنة نقوم بأخذ الفرق المناسب .

وبعد التحقق من توفر شرط سلكون السلسلة يتم تطبيق خطوات هذه الطريقة والتى تتلخص فى :

١. التعرف على النموذج Identification .
٢. تقدير النموذج Estimation .
٣. تشخيص النموذج Diagnostic .
٤. التنبؤ Prediction .

إن القراءات المسجلة زمنياً قد تأخذ نمطاً معيناً قد يكون تصاعدياً أو تنازلياً أو ثابتاً ، لذا فإن التنبؤ المستقبلى يتطلب أخذ هذا النمط بعين الاعتبار . كما أن استخدام عدد مناسب من القراءات يعطى نموذج مناسب يكون أكثر دقة خاصة عند استخدام أسلوب السلاسل الزمنية حيث يتم التنبؤ بالظاهرة مستقبلاً من البيانات المسجلة سابقاً دون الحاجة للحصول على بيانات

للمتغيرات الكثيرة التى تعتمد عليها الظاهرة ، ولصعوبة تمثيل وحصر هذه المتغيرات خاصة وأن هذه النماذج تعطى وزناً أكثر للقراءات القريبة ، ويقلل هذا الوزن كلما ابتعدنا عن القراءات المسجلة حالياً وهذا أكثر منطقية .

ونقوم بدراسة السلاسل الزمنية بهدف (والتر فاندل، ١٩٩٢) :

١. الحصول على وصف دقيق للملامح الخاصة للعملية التى تتولد منها السلسلة الزمنية .
٢. إنشاء نموذ جلتفسير وشرح سلوك السلسلة بدلالة متغيرات أخرى بربط القيم المشاهدة ببعض قواعد سلوك السلسلة .
٣. استخدام النتائج التى نحصل عليها للتنبؤ بسلوك السلسلة فى المستقبل وذلك اعتماداً على معلومات الماضى .

وسنقوم باستخدام تحليل السلاسل الزمنية باستخدام نموذج ARIMA ذو المتغير الواحد ويسمى هذا الأسلوب فى التحليل بأسلوب بوكس – جنكنز ويعتمد هذا الأسلوب على استخراج التغيرات المتوقعة للبيانات المشاهدة وتتجزء السلسلة إلى عدة مكونات أو عناصر تسمى ثلاثة مرشحات خطية هى مرشح السلكون Integrated Filter ومرشح الانحدار الذاتى Autoregressive Filter ومرشح المتوسطات المتحركة Moving Average Filter وتنتهى عملية تحليل البيانات باستخدام أسلوب بوكس – جنكنز عندما تبقى عناصر متناهية فى الصغر لا نستطيع الحصول منها على أية معلومات إضافية . والنموذج العام لـ ARIMA(P,D,Q) للبيانات الأصلية Z_t .

$$\Phi(\beta)wt = \delta + (\beta)a_t$$

حيث $\Phi(\beta) = I - \Phi_1(\beta) - \dots - \Phi_p\beta^p$

$$\theta(\beta) = I - \theta_1\beta - \dots - \theta_p\beta^p$$

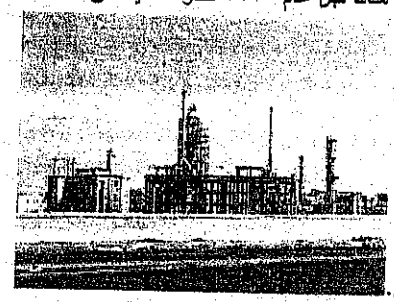
$$wt = \begin{cases} {}^{d}_{s}Z_t & d>0 \\ Z_t^d = 0 \end{cases}$$

الثابت δ ، متغيرات عشوائية a_t ، رتبة عملية الانحدار الذاتى P ، عدد الفروق D ، رتبة عملية المتوسطات المتحركة Q .

الفصل الثالث : الغاز الطبيعي

(٣-٢) استكشافه :

لم يكتشف الغاز الطبيعي بكميات تصلح للاستغلال التجارى إلا في عام ١٩٦٧ حين اكتُشِفَ حقل أبو ماضي في وسط الدلتا الذي كان بداية الاستكشافات الكبرى للغاز الطبيعي في مصر، وتبعه اكتشاف حقل أبى قير البحرى في البحر المتوسط في عام ١٩٦٩ وهو أول حقل بحري للغاز الطبيعي في مصر ثم حقل أبى الغراديق في الصحراء الغربية في عام ١٩٧١، وأدت النتائج المُشجِعة لتلك المرحلة المبكرة لتوسع عمليات البحث في الدلتا والصحراء الغربية وفي مياه البحر المتوسط التي بدأت الاستكشافات الأولية فيها عام ١٩٧٥، إلا إنه لم تبدأ حملات الاستكشاف المكثفة هناك قبل عام ١٩٩٥ لتقود للعديد من اكتشافات الغاز التجارية منذ عام



١٩٩٨ وحتى الآن.

(٣-٣) احتياطياته :

تمتلك مصر من الإحتياطى البترولى (المنتجات البترولية و" الغاز الطبيعي " والمتكثفات) ٣,٨ مليار برميل مكافئ عام ١٩٨٢/١٩٨١ ووصل بنهاية يونيه ٢٠٠٧ لحوالي ١٦,٩ مليار برميل مكافئ، ويمثل احتياطى الغاز الطبيعي حوالي ٧٥% من هذه الاحتياطيات.

وقادت الاكتشافات الكبيرة بالإضافة لتطوير حقول الإنتاج وجهود الاستكشاف المكثفة وتطبيق أحدث الطرق والتقنيات الحديثة بالإضافة للإنشاءات المتواصلة في الشبكة القومية للغازات الطبيعية لرفع الاحتياطيات المؤكدة من الغاز الطبيعي حيث زاد الاحتياطي المؤكد من الغاز الطبيعي من ٥,٤ تريليون قدم مكعب عام ١٩٧٩ إلى ٧٢,٣ تريليون قدم مكعب في عام ٢٠٠٦، بل إنه قفز في عام واحد من ٣٦ تريليون قدم مكعب إلى ٥١ تريليون قدم مكعب، وفى خلال السنوات ٢٠٠٢ - ٢٠٠٥ تم اكتشاف ٦ تريليون قدم مكعب من الاحتياطيات سنوياً، بجانب حوالى ١٠٠ تريليون قدم مكعب احتياطيات مرجحة لم تكتشف بعد طبقاً لتقديرات الشركات العاملة في مصر.وتمتلك مصر حوالى ١% من الاحتياطى العالمى، وتضم ثاني أكبر احتياطيات محتملة للغاز الطبيعي بالمياه العميقة في العالم بعد خليج المكسيك تصل لحوالى ٧٠ تريليون قدم مكعب، ويجرى العمل حالياً في خطة تنتهى عام ٢٠١٠ لإضافة ٣٠ تريليون قدم مكعب إلى احتياطيات الغاز الطبيعي في مصر باستثمارات ١٠ مليارات دولار. وطبقاً لأرقام عام

٢٠٠٥ فإن مصر هي الدولة رقم ١٨ بين ١٠٢ دولة لديها احتياطات مؤكدة من الغاز الطبيعي، وتضم حوالى ١% من الاحتياطى العالمى.

وهذ الاحتياطى الحالى يكفى مدة ٣٤ عاماً فقط للاستهلاك والتصدير طبقاً لمعدلات الاستهلاك الحالية، مما أدى لسعى الدولة للبحث عن بدائل أخرى للغاز الطبيعي خاصة في مجال توليد القوى الكهربية الذي يستهلك معظم الإنتاج المحلى خاصة مع تزايد استهلاك الطاقة الكهربية الذي ينمو بمعدل ٧,٣٢% خلال السنوات الأربع الأخيرة ٢٠٠٢ – ٢٠٠٥، فبدأ التفكير في الطاقة البديلة مثل طاقة الرياح والطاقة الشمسية، ولا تمثل الطاقة البديلة حالياً سوى ٣% فقط من استهلاك الكهرباء في مصر وإن كان من المخطط زيادة النسبة لتصل إلى ١٣% في عام ٢٠١٠ .

الجزء التطبيقي : ويتضمن الفصول التالية :

أظهرت الدراسة في هذا المبحث اختبار السببية أن هناك العلاقة سببية بين الانتاج والاستهلاك للغاز الطبيعي في اتجاه واحد ولوغاريتم الاستثمار الأجنبي المباشر و لوغاريتم عدد السكان إلي لوغاريتم الناتج المحلي ، وعليه يعد الانتاج والاستهلاك متغيرات تفسيرية مستقلة .
وتلاحظ التقارب الشديد بين معايير جودة النموذج لكل من الدالة التحويلية والشبكات العصبية و إن كانت الأفضلية لدالة التحويل خاصة مع السلاسل الزمنية الطويلة .

المبحث الثاني : الشبكات العصبية :

ولمعرفة الأسلوب الأكثر ملائمة للتوفيق للبيانات باستخدام المعايير الإحصائية لقياس قدرة النموذج علي التنبؤ وللمفاضلة بين الأساليب تم استخدام المعايير التالية بجدول (٢٧):

١. معامل التحديد Coefficient of Determination(R^2)
٢. المتوسط النسبي لخطأ التنبؤ المطلق (MAPE)
٣. الجذر التربيعي لخطأ التنبؤ Root Mean Square Of Error(RMSE)
٤. متوسط القيمة المطلقة للأخطاء (MAE)
٥. معامل ثيل Theil Coefficient (T.C)

جدول (٢٧)

النموذج	الجذر التربيعي لخطأ التنبؤ(RMSE)	المتوسط النسبي لخطأ التنبؤ المطلق (MAPE)	متوسط القيمة المطلقة للأخطاء (MAE)	معامل ثيل	معامل التحديد
الانحدار المتعدد	٠,٩١٤	٤,٢٠٠	0.746	٠,٠٢٥	٠,٦٦٦
السلاسل الزمنية	٠,٤٩٩	٢,٢٨٥	0.417	٠,٠١٣	٠,٨٨٥
الشبكات العصبية	٠,٦٥٢	٢,٨٨٦	0.515	٠,٠١٨	٠,٧٦٤

هذا وقد أظهرت النتائج افضلية نموذج دالة التحويل على النماذج الأخرى فقد تحقق معايير جودة النموذج به من خلال نتائج الجدول السابق وأيضا من خلال استقرار البواقى وعدم وجود مشكلة الارتباط الذاتى التسلسلى للبواقى الى جانب عدم وجود مشكلة فى تباينات الخطأ العشوائى الى جانب اعتدالية البواقى.

وبدراسة الفروق الاحصائية بين الانتاج والاستهلاك من الغاز الطبيعى فقد استخدمت الباحثة اختبار ت لعينتين مرتبطتين وكانت النتائج كما هو موضح بالجدول التالى (٢٨)

جدول (٢٨)الفروق الاحصائية بين الانتاج و الاستهلاك

المتغيرات	المتوسط	الانحراف المعيارى	اختبار ت	درجة الحرية	المعنوية
الاستهلاك	28.31	15.372	-3.365	25	0.002
الانتاج	32.88	20.491			

الفصل الخامس : النتائج والتوصيات

و يعرض هذا الفصل النتائج و التوصيات التي يمكن تلخيصها كما يلي :

١ – أن الدالة التحويلية و الشبكات العصبية أظهرت تميزا ملحوظا عن الاساليب التقليدية " الانحدار المتعدد التدريجي " ٠

٢ – أظهرت النتائج وجود أثر معنوي علي الناتج المحلي الاجمالي باستخدام نموذج دالة التحويل وقد بلغت القدرة التفسيرية ٠,٩٨

٣– ربط البحوث العلمية بأهداف السياسة السكانية لحل مشاكل التطبيق٠

٤– يجب علي كل من يقوم بدراسة يتطلب فيها نظرة مستقبلية أن يقوم باستخدام الشبكات العصبية الاصطناعية و أن يتم تحليلها باستخدام الاساليب الاحصائية الحديثة ، وذلك لتحقيق الاستفادة القصوي منها حيث أن الشبكات العصبية لديها السرعة و الدقة واستخدامها لا يتطلب شروطا كما في الاساليب الاحصائية ٠

٥ – وجد أن نتائج البرامج المتخصصة في الشبكات العصبية الاصطناعية أكثر دقة من الاساليب الاحصائية وهذا يتوقف علي المعايير المختارة حسب نوع النموذج المختار ٠

أولاً: المراجع العربية:

١- البدراني، ظافر مطر، الدباني، عمر اسماعيل (٢٠١٣م): "توفيق نموذج ديناميكي لعملية تصفية المياه في مدينة الموصل"، المجلة العراقية للعلوم الاحصائية، (٢٣)، العراق.

٢- البدراني، ظافر مطر، عبدالكريم حسين (٢٠٠٦م): "النماذج الحركية لدالة التحويل وتعدد المدخلات"، المجلة العراقية للعلوم الإحصائية، (٩)، العراق.

٣- بدري، مسعود والمطوع، أحمد وهادي، عقيل (١٩٩٦م): "استخدام تحليل التمايز والشبكات العصبية في التنبؤ بدرجة اعتمادية العميل المصرفي"، المجلة العربية للعلم الإدارية، الكويت، مجلد (٣)، عدد (٢)، ص [٢٩٥-٣١٥].

٤- بري، عدنان ماجد (٢٠٠٢م): "طرق التنبؤ الاحصائي"، الجزء الأول، جامعة الملك سعود، كلية العلوم، قسم الاحصاء وبحوث العمليات.

٥- البشير، زين العابدين عبدالرحيم (٢٠١٣م): "السلاسل الزمنية في اتجاهي الزمن والتكرار"، دار الجنان، السودان.

٦- توفيق، أماني مصطفي (٢٠٠٦م). "إدارة أخطار التلوث الناشئ عن صناعة البترول في جمهورية مصر العربية"، رسالة دكتوراه في التأمين، جامعة المنصورة.

٧- حاجي، جعفر، والمجيد، محمد عبدالهادي (١٩٩٩م): "الشبكة العصبية: التنبؤ بأسعار صرف الدينار الكويتي مقابل الدولار الأمريكي"، المجلة العربية للعلوم الإدارية، الكويت، مجلد (٦)، العدد (١)، ص [١٧-٣٥].

٨- حياوي، هيام عبدالمجيد، إسماعيل، أحمد سالم (٢٠١٢م): "توظيف نهج التحليل العاملي للتنبؤ بنماذج دالة التحويل"، المجلة العراقية للعلوم الاحصائية، (٢١)، العراق.

٩- حياوي، هيام عبدالمجيد، محمد، شيرين تركي (٢٠١٢م). "مقارنة التنبؤ باستخدام نماذج دالة التحويل ونماذج الأنماط المتماثلة المضببة مع التطبيق"، المجلة العراقية للعلوم الاحصائية، (٤٢)، العراق.

١٠- خليل، شهباء إبراهيم (٢٠١١م): "تضمين استخلاص العلامة المائية في الملفات الفيديوية باستخدام التقنيات الذكائية"، مجلة الرافدين لعلوم الحاسب والرياضيات، المجلد ٨، العدد١، كلية علوم الحاسوب والرياضيات، جامعة الموصل، ص [٢٥-٤٤]

١١- الخياط، باسل، يونس، ذكي، عزة، حازم (٢٠٠٥م): "استخدام الشبكات العصبية في التكهن بالسلاسل الزمنية، المجلة العراقية للعلوم الإحصائية، (٨)، العراق.

١٢- ديفيد، م. سكا بورا (٢٠٠٢م): "بناء الشبكات العصبية"، ترجمة: فهد بن عبدالله التركي، الرياض، جامعة الملك سعود.

١٣- زريفة، بديع (٢٠٠٩م): "استخدام الشبكة العصبونية للتنبؤ بمقدار الطلب على الطاقة الكهربائية"، مجلة جامعة دمشق للعلوم الهندسية، المجلد ٢٥، العدد ٢، كلية الهندسة الميكانيكية والكهربائية، جامعة دمشق، ص [١٧٩-٢٠١].

١٤- سولاف صبري، جواد كاظم (٢٠١٢م): "التنبؤ بمبيعات البنك المركزي من العملة الأجنبية"، المجلة العراقية للعلوم الاقتصادية، العراق.

١٥- الصفار، عبدالكريم (١٩٩٨م): "استخدام الشبكات العصبية للتنبؤ باستهلاك الكهرباء في الكويت"، مجلة العلوم الاقتصادية والإدارية، الامارات العربية المتحدة، العدد (١٤)، ص [١٢٥-١٤٣].

١٦- طاهر، جميل (١٩٩٧م): "النفط والتنمية المستديمة في الأقطار العربية: الفرص والتحديات".

١٧- العباسي، عبدالحميد محمد (٢٠٠٤م): "المقارنة بين أسلوب الشبكات العصبية والمساريما للتنبؤ بأعداد الوفيات الناتجة عن حوادث المرور بالكويت"، المجلة العربية للعلوم الإدارية، الكويت، مجلد (١١)، العدد (٣)، ص [٣٣٣-٣٥٩].

١٨- العباسي، عبدالحميد محمد (٢٠٠٥م): "المقارنة بين أسلوب الشبكات العصبية والأساليب الإحصائية التقليدية للتنبؤ بأعداد الوفيات الناتجة عن حوادث المرور بالكويت"، المجلة المصرية للسكان وتنظيم الأسرة، معهد الدراسات والبحوث الإحصائية، جامعة القاهرة، مصر.

١٩- العباسي، عبدالحميد محمد (٢٠٠٩م): "التحليل المتعدد المتغيرات باستخدام SPSS"، معهد الدراسات والبحوث الاحصائية، جامعة القاهرة، مصر.

٢٠- العباسي، عبدالحميد محمد (٢٠١٠): "التحليل الاحصائي المتقدم باستخدام SPSS"، معهد الدراسات والبحوث الاحصائية، جامعة القاهرة، مصر.

٢١- العبيدي، محمود خليل إبراهيم (٢٠٠٠م): "الشبكات العصبية الاصطناعية"، مجلة أبحاث الحاسوب، المجلد (٤)، العدد (١)، الجامعة التكنولوجية، بغداد.

٢٢- العتيبي، فوزي (٢٠٠٣م): "استخدام السلاسل الزمنية والشبكات العصبية في التنبؤ بالأرقام القياسية"، رسالة ماجستير في الإحصاء التطبيقي، جامعة قناة السويس، بورسعيد.

٢٣- عماد يعقوب (٢٠٠٩م): "استخدام نماذج بوكس جنكينز ونماذج الشبكات العصبية للتنبؤ في السلاسل الزمنية الاقتصادية"، أطروحة دكتوراه غير منشورة، جامعة السودان للعلوم والتكنولوجيا، السودان.

٢٤- العواد، مساعد ناصر جاسم (٢٠٠٨م): "نظرة تحليلية للأهمية الاقتصادية للبترول والغاز الطبيعي في منطقة الشرق الأوسط"، قسم هندسة البترول والغاز الطبيعي، كلية الهندسة، جامعة الملك سعود، الرياض، المملكة العربية السعودية.

ثانيا: المراجع الأجنبية:

1- A. Yaffee, Robert: "Introduction to Time Series Analysis and Forecasting with Application of Sea an SPSS", Stat University of New York Health Science.
2- Abraham, B., & Ledolter, J. (1983): "Statistical Methods for Forecasting", New York, Wiley.
3- Shankar . T . N .(2008) . " Neural Networks " . UNIVERSITY SCIENCE PRESS .
4- Al Neamy, Jamal Salahaldeen (2006): "Brain Tumors Images Diagnosis Using Hybrid Intelligence Techniques", Ph. D. Thesis, the College of Computers and Mathematics Science, University of Mosul.
5- Allen Rubin (2010): "Evidence-Based Practice and Evaluation", 2nd Edition, ISBN-01: 0495602892.
6- Anderson, O. D. (1976): "Time Series Analysis and Forecasting", London, Butterworths.
7- Aneiros-Perez, G., Cao, R., & Vilar-Fernandez, J. M. (2010): "Functional Methods for Time Series Prediction: A Nonparametric Approach", Journal of Forecasting, PP. 3-6. doi: 10.1002/for.1169.
8- B. Yegnanarayanan (1999): "Artificial Neural Network", New Delhi, Prentice-Hall of India.
9- Beck, Nathaniel & Jonathan N. Katz (1995): "What to Do (and Not to Do) with Time-Series Cross-Section Data", American Political Science Review, 89, 634-647.
10- Bisgaard, S. & Kulahci, M. (2011): "Time Series Analysis and Forecasting By Example", Published by John Wiley & Sons, Inc., Hobken, New Jersey.
11- Bobko, Philip (2001): "Correlation and Regression", 2nd Edition, Thousand Oaks, CA: Sage Publication, Introductory text with Includes coverage of range restriction, trivariate correlation.
12- Bowerman, Bruce L. & Richard T. O'Connell (1979): "Time Series and Forecasting: An Applied Approach", Duxbury Press.
13- Box, George E. P. & Gwilym M. Jenkins (1976): "Time Serise Analysis: Forecasting and Control", Revised Edition, Holden-Day.
14- Chatfield, C. (1989): "The Analysis of Time Series", 4th Edition, Chapman and H London.
15- Cleveland, W. S. (1979): "Robust Locally Weighted Regression and Smoothing Scatterplots", Journal of the American Statistical Association, 78 (368), 892-836. doi: 10.1080/ 01621459.1979.10481038.
16- Coughlin, K. & Eto, J. H. (2010): "Analysis of Wind Power and Load Data at Multiple Time Scales", Lawrence Berkeley National Laboratory, University of California, Berkeley, California, PP. 3-5.
17- Cryer, J. D., & Chan, K. S. (2008): "Time Series Analysis with Application in R", 2nd Edition, Springer, New York.
18- David P. Doane, Lori E. Seward (2011): "Measuring Skewness: A Forgotten Statistic?", Oakland University, Journal of Statistics Education Volume 19, Number 2.
19- Donald J. Wheeler (2012): "Problems with Skewness and Kurtosis Part Tow", American Statistical Association and the American Society for Quality.
20- Douglas C. Montgomery (2008): "Introduction to Time Series Analysis and Forecasting".