استخدام الذكاء الاصطناعي والانحدار الحصين للتنبؤ بمعدلات المواليد في مصر

د/ عبدالوهاب السيد حجاج
أستاذ الإحصاء المساعد ورئيس قسم الإحصاء ووكيل كلية التجارة بنين – جامعة الأزهر

د/ محمد عبدالقادر
مدرس الإحصاء
كلية التجارة بنين – جامعة الأزهر

ملخص البحث

نشأت مشكلة هذا البحث من الناحية السكانية حين عجزت العديد من الدراسات عن التنبؤ بمعدلات المواليد ومعدلات الإنجاب بشكل أقرب للواقع، فقد اعتمدت الإسقاطات السكانية علي استخدام دالة النمو الأسي وعلي توفيق منحني سلسلة زمنية من عدد من السكان في فترات سابقة، كما أعتمد مشروع مصر ٢٠٢٠ علي فرضية خطية معدل الإنجاب بناء علي خطية العلاقة ، ومن الناحية الإحصائية فإن وجود القِّيم الشاذة تؤثر في توزُّع بواقي النموذج وسيكون توزُّيعا ملتوٌيا وعلُّيه سيكون مخالفا لشروط طرُّيقة المربعات الصغرى المعتادة ، لذلك تم البحث عن طرائق بدُّيلة للتقدُّر. ومن ثم استخدام الشبكات العصبية كأسلوب من أساليب الذكاء الاصطناعي بجانب أسلوب الانحدار الحصين الشائع استخدامه في حالة البيانات الشاذة اوالمتطرفة ،ولدراسة التطبيقية تم اختيار عينة كبيرة حجمها ٦٠ مفردة، حيث تمتد السلسة الزمنية منذ عام ١٩٦٠ حتى عام ٢٠١٩. ومن خلال الاستفادة من خوارزميات برنامج MATLAB وبرنامج R تم حساب القيم التنبؤية لقيم معدلات المواليد، وتم حساب دقة هذه التنبؤات وذلك بواسطة مقاييس إحصائية تعتمد على الفرق بين القيمة التنبؤية و القيمة الفعلية (البواقي) وهى (MSE,MAD)، ولقد تم الحصول على نتائج (MSE, MAD) لتنبؤات الشبكات العصبية NN)) وتنبؤات الانحدار الحصين Rob-Re)) واتضح أن جميع قيم MAD,MSE للشبكات العصبية الاصطناعية كانت اقل من أسلوب الانحدار الحصين ، كذلك كان المتوسط العام لمربعات الأخطاء للشبكات

MSE(NN)=2.944 اقل من المتوسط العام للانحدار الحصين MEE(Rob-Re)=9.015، مما يعنى أفضلية أسلوب الشبكات العصبية الاصطناعية عن الانحدار الحصين للتنبؤ بمعدلات المواليد في ظل وجود القيم الشاذة اوالمتطرفة.

مقدمة:

لقد أصبحت دراسة السكان محوراً رئيساً، وتشتق منه كثير من الدراسات في علوم مختلفة، وبخاصة بعد أن شهدت الزيادة السكنية في المرحلة الأخيرة -في معظم دول العالم -طفرة كبيرة وتعد دراسة عدد السكان ونموهم خطوة أساسية ومهمة لعمليات التخطيط والتنمية الاقتصادية والاجتماعية والعمرانية،كما تكمن أهمية دراسة حجم السكان ومكوناته، ليس فقط لمعرفة عدد السكان الحالي ونموه في الفترة السابقة، بل في تحديد التزايد السكاني في المستقبل، وبالتالي تقدير عدد السكان في السنوات المقبلة.) باحاج).

وفي ظل ثورة المعلومات وما نتج عنها من تدفق هائل للبيانات أصبحت هناك حاجه ملحة لتوافر قواعد بيانات دقيقة وممثلة لواقع الظاهرة أو الظواهر محل الدراسة ثم استخدام أسلوب علمي وأدوات تحليل مبتكرة لإجراء معالجة الكترونية Electronic processing بدلا من المعالجة اليدوية Manual processing لهذه البيانات. والهدف هو الحصول على معلومات كافية كما وكيفا لاحتياجات الباحثين وصانعي القرارات. ومع تطور الحاسب الآلي - والذي أصبح عاملا رئيسا في كافة أنشطة الحياة تقريبًا - توصل الباحثون لنوع من البرامج تعتمد طريقة بنائها على تقنية تحاكي الإنسان في طريقة اتخاذه لقراراته اليومية تسمى الذكاء الاصطناعي.

ويعرف الذكاء الاصطناعي Artificial intelligence بأنه: علم يتناول كيفية جعل الآلة – أي الحاسب – تؤدى عمليات مناظرة لقدرات البشر العقلية. وقد ظهر الذكاء الاصطناعي في الخمسينات من القرن الماضي نتيجة الثورة التي حدثت في مجالي المعلومات والتحكم الآلي ، وترمى أبحاثه إلى تحقيق هدفين رئيسين: الأول: الوصول إلى فهم عميق للذكاء الإنساني عن طريق محاكاته. الثاني: الاستثمار الأفضل للحاسب الآلي والعمل على استغلال إمكاناته كافة ، وخصوصا بعد التطور السريع في قدرات الحاسبات وانخفاض ثمنها. وتعد الشبكات العصبية الاصطناعية (Network ANN)

Artificial Neural من أهم الطرق المرنة التي لديها القدرة على التعلم الذاتي السريع وسرعة التغير في البيئة الديناميكية ، وذلك من خلال قواعدها الخاصة ووفق منهجية محددة وتعزيز هذا المنهج عبر التدريب بطريقة تحاكى فيها عمل الخلايا العصبية الدماغية من حيث بنيتها ومعالجتها للمدخلات.

(١-٢) مشكلة البحث

من الناحية السكانية: نشأت مشكلة البحث حين عجزت العديد من الدراسات عن التنبؤ بمعدلات المواليد ومعدلات الإنجاب بشكل أقرب للواقع، فقد اعتمدت الإسقاطات السكانية علي استخدام دالة النمو الأسي وعلي توفيق منحني سلسلة زمنية من عدد من السكان في فترات سابقة، كما أعتمد مشروع مصر ٢٠٢٠ علي فرضية خطية معدل الإنجاب بناء علي خطية العلاقة خلال العقدين السابقين علي الدراسة، وعلي افتراضات نظرية التحول الديموجرافي.

من الناحية الإحصائية: أن وجود القُيم الشاذة تؤثر في توزُّع بواقي النموذج وسيكون توزُّيعا ملتوُّيا وعلُّيه سيكون مخالفا لشروط طرُّيقة المربعات الصغرى المعتادة لذلك يتم البحث عن طرائق بدُّيلة للتقدُّر.

(١-٣) هدف البحث

يهدف البحث إلى اقتراح نموذج ملائم للتنبؤ بمعدلات المواليد، من خلال المقارنة بين:

- نموذج الانحدار اللامعلمي الحصين.
- الشبكات العصبية كأسلوب من أساليب الذكاء الاصطناعي.

(١-٤) أهمية البحث

عمدت الكثير من الأبحاث إلى التنبؤ بمعدلات المواليد لما له من أهمية في التحكم بهذه الظاهرة، والتخطيط لمجابهة تبعاتها، تعليميا وصحيا و علي مستوي الرعاية الاجتماعية، فضلا عن حسن استغلال الهبة الديموجرافية المتوقعة.

وبالرغم من أن فرصة استغلال النافذة الديموجرافية مطلع هذا القرن قد أفلتت، إلا أن الأمل معقود علي حسن استغلال الفرصة القادمة والمرتهنة باستمرار السياسة

السكانية الحالية الرامية لخفض معدلات المواليد والإنجاب، فضلا عن توافر تنبؤات علمية دقيقة باستخدام طرق تنبؤ حديثة. ومن الممكن تلخيص أهمية الدراسة في اتجاهين هما:

١- **أهمية تطبيقية**: وتتركز الأهمية التطبيقية لهذه الدراسة في أنها تساعد المؤسسات الكبيرة والدول في وضع خططها المستقبلية بناء على قراءة لتطور أحوال السكان ومعدلات مواليدهم.

٢- **أهمية إحصائية**: تكمن الأهمية الإحصائية لهذه الدراسة في استخدام طرائق وأساليب إحصائية متطورة مثل الشبكات العصبية الاصطناعية كأحد طرق الذكاء الاصطناعي للتنبؤ بمعدلات المواليد على المدى البعيد باستخدام بيانات لمدة زمنية طويلة.

(١-٥) الدراسات السابقة:

لقد قامت العديد من الأبحاث بدراسة التنبؤ بمعدلات المواليد بطرق إحصائية عديدة ولكن قليل من الدراسات هي التي تناولت استخدام الذكاء الاصطناعي:

فلقد تم استخدام السلاسل الزمنية للتنبؤ بالمواليد عن طريق ($saboia1977$)،

($Mcdonald\ 79,81$) حددا شكل العلاقة وذلك بالتطبيق علي بيانات مواليد استراليا، وقام ($Gandon1980$) بدراسته عن استخدام السلاسل الزمنية في التنبؤ بمواليد لندن.

ثم قدم ($Land1986$) دراسته حول نمذجة العمليات الديموجرافية باستخدام السلاسل الزمنية.

كما قام ($de-beer1988$) دراسته حول المقارنة بين التمهيد الأسي ونماذج $ARIMA$ في التنبؤ بمعدلات المواليد.

ثم استخدم ($Carter1996$) أسلوب المقارنة بين $ARIMA$ ونماذج $Structural-time-series$ في التنبؤ بعدد الوفيات بالولايات المتحدة.

كما استخدم ($jose1998$) السلاسل الزمنية في التنبؤ بمواليد البرازيل.

وقد شهد عام ٢٠١٣ صدور عدة أبحاث حول استخدام السلاسل الزمنية في التنبؤ بعدد المواليد إذ قدم (*Bravo* 2013) دراسته عن استخدم نماذج *ARIMA* في التنبؤ بالمواليد والوفيات لسكان ايطاليا.

ودراسة (مصباح ٢٠١٣) في التنبؤ بمواليد عطبرة بالسودان. ثم (*Onwuka* 2013) بدراسته عن استخدام السلاسل الزمنية في دراسة مواليد الشمال الغربي النيجيري.

ثم (*Essuman* 2017) الذي استخدم سلاسل بوكس وجينكينز في التنبؤ بمواليد غانا.

كما قام قارن (*Bravo* 2019) بين السلاسل الزمنية الخطية واللاخطية في التنبؤ بالمواليد.

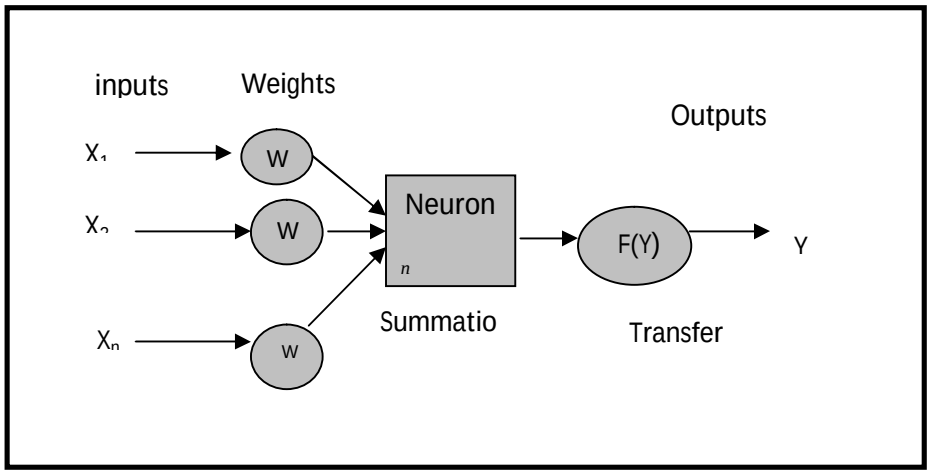
ثم (*Senyefia* 2019) استخدام السلاسل الزمنية في التنبؤ بمواليد غانا.

وبالنسبة لهذا البحث الذي يدرس معدلات المواليد للسكان في مصر فقد تم اختيار عينة كبيرة حجمها ٦٠ مفردة، حيث تمتد السلسة الزمنية منذ عام ١٩٦٠ حتى عام ٢٠١٩. وقد تم اختيار عام ١٩٦٠ كبداية للسلسلة لأنه بداية فترة الستينات التي شهدت بداية انخفاض معدلات المواليد في مصر تحت حاجز ٥٠ مولود لكل ألف من السكان. كما تم استخدام معدل المواليد وليس عدد المواليد وذلك لارتباط عدد المواليد بعدد السكان.

ويتضح مما سبق انه على الرغم من كثرة الدراسات التى قدمت للتنبؤ بمعدلات المواليد إلا انها لم تستخدم الذكاء الاصطناعي كطريقة جديدة ومتطورة.

(١-٦) الشبكات العصبية الاصطناعية

تُعد الشبكات العصبية الاصطناعية Neural Networks (ANN) Artificialأحد أهم طرق الذكاء الاصطناعي، و تتمحور فكرتها حول محاكاة قدرة العقل البشرى على التعرف على الأنماط ، وتمييز الأشياء ، باستخدام الحاسب الآلي من خلال إتباع عملية التعليم الذاتية التي تحدث في العقل ، والتي يتم فيها الاستفادة من الخبرات السابقة في سبيل الوصول إلى أفضل نتائج في المستقبل.(نظر: حاجى، و المحميد ، (١٩٩٩) ،ص ١٩).



شكل (١) نموذج لوحدة تشغيل

والشكل(١) السابق يبين وحدة تشغيل بسيطة داخل الشبكة العصبية (انظر: Turban, et al. (2005))، حيث يكون لكل وحدة تشغيل مسار إدخال واحد أو أكثر تنحصر مهمته في نقل المعلومات من العالم الخارجي إلى وحدة التشغيل الذي تقوم بدورها بعملية التجميع البسيطة ، ثم تحويل المعلومات بواسطة دالة تفعيلية تعرف بدالة التحويل ، بعد ذلك تنقل المعلومات كمخرجات من خلال مسار المخرجات.

وبشكل رياضي فان كل خلية عصبية تستقبل عدد من إشارات المدخلات Input (Signals $(X_1, X_2, \ldots X_n$ والتي تماثل التفرعات العصبية Dendrites ، ثم يتم ضرب كل المدخلات في الأوزان $(W_1, W_2, \ldots\ldots, W_n)$ التي تعد الوسيلة الأساسية لذاكرة الشبكات العصبية الاصطناعية طويلة المدى وتعبر عن درجة أهمية المدخلات، ، وحاصل الجمع summations يتم معالجته باستخدام دالة تحويل F(Y) Transfer function لنحصل منها على إشارات المخرجات Y)Output Signals).

(١-٦-١) تعليم الشبكة العصبية Learning of the Neural Network

تعلم الشبكة العصبية الاصطناعية هي العملية التي من خلالها تقوم الشبكة العصبية بتعديل نفسها استجابة للمدخلات من أجل الحصول على المخرجات المستهدفة. وبأسلوب آخر فهي عملية اكتساب المعرفة ، حيث تبحث الشبكة العصبية عن المعرفة من مجموعة بيانات العينة. وأثناء عملية التعلم تعدل الشبكة أوزان الاتصال على أساس

المدخلات الواردة حتى تقترب المخرجات من المخرجات الحقيقية أو المستهدفة (انظر Zurada (1992)).

ويتم تعليم أو تدريب الشبكة العصبية على اكتشاف العلاقات النمطية في البيانات بإتباع إحدى الطرق الآتية:

١- التعليم الإشرافي:Supervised Learning

تعد هذه الطريقة هي الأكثر انتشارا في تدريب الشبكة العصبية ، وفيها تجمع عينات من البيانات تحتوى قيماً للمتغيرات المدخلة وقيماً للمتغيرات المخرجة ،و تقوم الشبكة بمقارنة النتائج التي تقدرها للمتغيرات الخارجة لكل عينة مدخلة بالقيم الفعلية لهذه المتغيرات ، وبناء على ذلك تقوم الشبكة بإجراء التعديلات على أوزان الاتصال بهدف تقليل الأخطاء في النتائج ، ثم تعاد عملية التدريب عدة مرات إلى أن يتم الوصول إلى نتائج مقبولة.

٢- التعليم غير الإشرافي Unsupervised Learning

تتشابه هذه الطريقة مع طريقة التعليم الإشرافية إلا أنها تختلف عنها في أن العينات المستخدمة في عملية التدريب لا تتضمن أية قيم للمتغيرات الخارجة ، وتتكون البيانات الداخلة إلى الشبكة من عدة قطاعات أو مجموعات ، حيث تتدرب الشبكة في هذه الحالة على اكتشاف المميزات غير الظاهرة في مجموعة البيانات المستخدمة في عملية التدريب ، ومن ثم استخدام تلك المميزات في تقسيم بيانات المدخلات إلى مجموعات مختلفة فيما بينها ومتقاربة داخل كل مجموعة.

٣- التعليم بإعادة التدعيم: Reinforcement Learning

هذه الطريق خليط بين الطرقتين السابقتين ، حيث لا يفصح للشبكة العصبية عن القيم الحقيقية للمخرجات كما هو الحال في طريقة التدريب غير الإشرافية ولكن يشار للشبكة بصحة نتائجها المحصلة أو خطئها كما في طريقة التعليم الإشرافية.

(١-٦-٢) الهيكل العام للشبكة Typical Architectures

عملية تنظيم الخلايا العصبية في طبقات وكيفية الاتصال بين هذه الخلايا لتكوين الشبكة تسمى بهيكل الشبكة Architectures ، و بوجه عام من الممكن تقسيم هيكل الشبكة العصبية الاصطناعية إلى ثلاثة أنواع رئيسية هي:(١) شبكة وحيدة

الطبقة ذات التغذية الأمامية 2) , Single-Layer Feed forward Network) شبكة متعددة الطبقات ذات التغذية الأمامية Multi-Layer Feed forward 3) , Network) شبكة متعددة الطبقات ذات التغذية المرتدة Multi-Layer Recurrent Network.(انظر: Demuth & Beale 1998) و سنوضح كل نوع فيما يلي:

١- شبكة وحيدة الطبقة ذات التغذية الأمامية

وهى أبسط هياكل الشبكات العصبية الاصطناعية، والنوع الأكثر شيوعا منه يسمى Perceptron ، وهو ابسط صيغة للشبكات العصبية الاصطناعية،كما أنه النموذج الذي تبنى عليه الأنواع الأخرى من الشبكات وحيدة الطبقة، كما يعد ابسط أنواع الشبكات العصبية أمامية التغذية حيث تنتقل المعلومات من الطبقة المدخلات إلى طبقة المخرجات مباشرة ، كما أنها تتعلم عن طريق التعليم الإشرافي. والشكل (١) السابق يوضح هيكل شبكة عصبية وحيدة الطبقة بتغذية أمامية. ومن خلال تلك الشبكة تتم عملية التعلم عن طريق قيام وحدات المعالجة بعدة خطوات رئيسية هي:

الخطوة الأولى: التحديد العشوائي للقيم الابتدائية للأوزان $W_1, W_2, \ldots, W_n$ ، وللقيمة الحدية θ (Threshold Value) وذلك في المدى [-٠.٥,٠.٥].

الخطوة الثانية:عملية التنشيط: فبعد أن تستقبل كل وحدة معالجة (نيرون) Neuron العديد من الإشارات المدخلة، و حساب المجموع الموزون لهذه المدخلات المستخدمة و الذي عادة ما يتم باستخدام دالة المجموع Summation Function الآتية:

$$X = \sum_{i=1}^{n} w_i x_i \qquad \textbf{(1)}$$

X: صافى المدخلات الموزونة للخلية العصبية

W_i: الأوزان النسبية لعقد الاتصال ما بين الطبقات

x_i: قيمة المدخلات i

n: عدد مدخلات الخلية العصبية

يتم تنشيط المدرك Perceptron عن طريق تطبيق المدخلات $X_1, X_2, \ldots X_n$ ، والمخرجات المرغوب بها $T(k)$ ، وحساب المخرجات الفعلية عند التكرار k كما يلي:

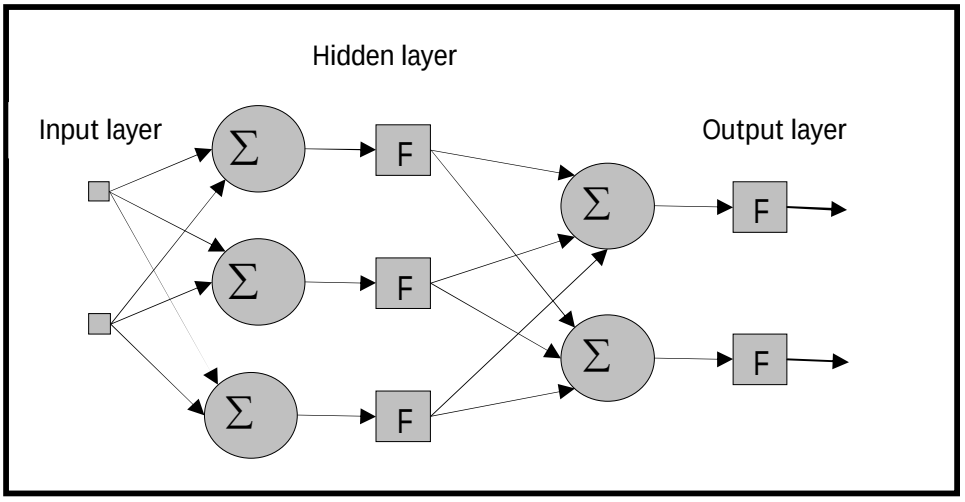
$$y(k) = step[\sum_{i=1}^{n} X_i(k)W_i(k) - \theta] \qquad \textbf{(2)}$$

وهذا النوع من دوال التحويل(التنشيط) يسمى بدالة الخطوة Step Function.

وتجدر الإشارة إلى انه بجانب دالة الخطوة يوجد كثير من دوال التنشيط ولكن وجد أن قلة فقط منها لها تطبيقات عملية وفى أغلب الأحوال دالة التحويل تجمع من توليفة خطية (Al-Shawadfi 2003) .

متعددة الطبقات ذات التغذية الأمامية

تتكون الشبكة متعددة الطبقات من واحد أو أكثر من الطبقات (أو المستويات) من عقد الاتصال Nodesتسمى طبقات مخفية Hidden Layers توجد بين طبقات المدخلات Input Layers وطبقات المخرجات Output Layers ، وفى الشبكات ذات التغذية الأمامية تنتشر إشارات المدخلات في الاتجاه الأمامي من طبقة لطبقة، ولا يسمح بالرجوع للخلف، ويمكن لهذا النوع من الشبكات حل العديد من المشاكل المعقدة التي لا تستطيع الشبكة ذات الطبقة الواحدة حلها ، ولكن تدريبها قد يستغرق وقتا أطول.انظر: Hijazi (2007)



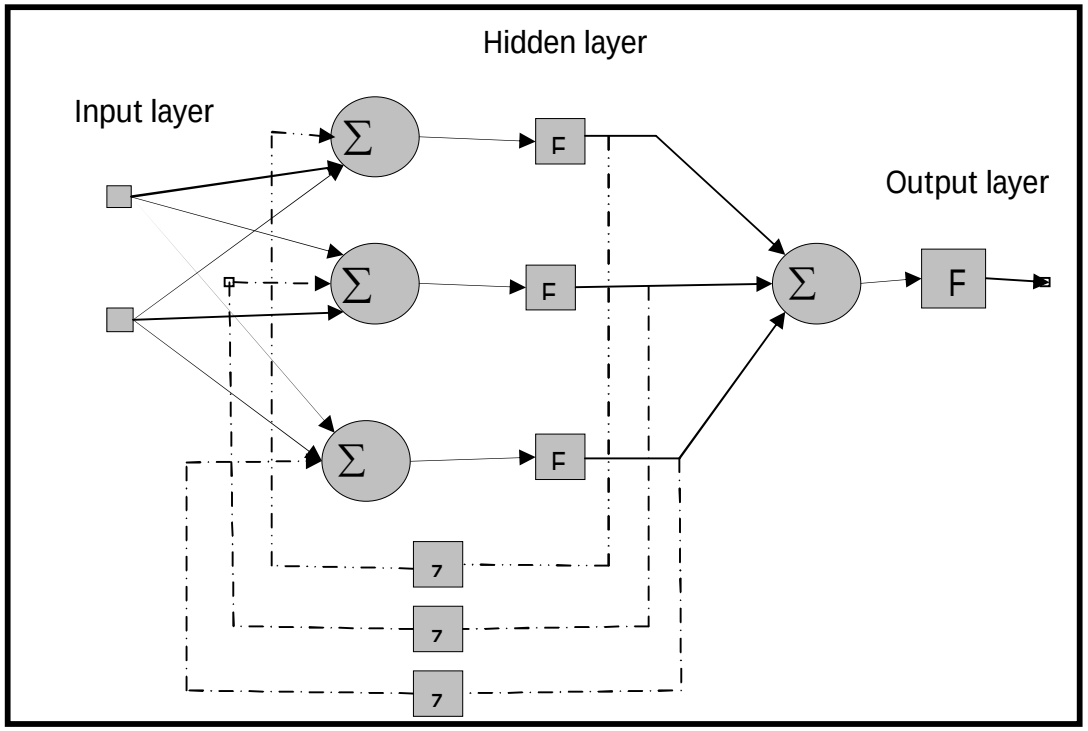
شكل(٢) الشبكة متعددة الطبقات ذات تغذية أمامية

وتقوم العناصر الحسابية في الطبقة المخفية بإجراء عملية الجمع الموزون وترسل إشارة مخرجات (أنماط تنبيه) إلى طبقة المخرجات التي تقوم هي الأخرى بعملية الجمع الموزون وتحديد أنماط المخرجات للشبكة كلها.(انظر:(٢٠٠٤) نجينفيتسكى). كما انه بزيادة عدد وحدات المعالجة داخل الطبقة الخفية أو إضافة أكثر من طبقة مخفية في الشبكة ، يسمح للشبكة بالتعامل مع كثير من الدوال المعقدة والتقريب الدقيق للدوال المستمرة وهذا ما نصت عليه نظرية (Cybenko). (أنظر Cybenko (1989)).

أما بالنسبة لعملية التعليم فإنها تستمر بنفس الطريقة في الشبكة وحيدة الطبقة ، حيث تحّدث الأوزان باستخدام التعلم بالتغذية الخلفية Back propagation Learning ، وقبل هذا التحديث يوجد خطوتين متتاليتين أولهما أماميةForward والأخرى تراجعية Backward ، ففي الخطوة الأمامية يتم حساب مخرجات الشبكة من البيانات المدخلة ، ومقارنة هذه المخرجات مع المخرجات المستهدفة عن طرق حساب الخطأ. أما في الخطوة التراجعية فتقوم الشبكة بتعديل الأوزان النسبية بهدف تصغير الخطأ ، وتسمى العملية التي تضم المرحلتين السابقتين بالدورة (Epoch)، ويتم تكرار هذه الدورة إلى أن نصل لأقل مجموع مربعات للخطأ.

٣- شبكة متعددة الطبقات ذات التغذية المرتدة

وهى النوع الثاني من الشبكات العصبية متعددة الطبقات ، و هذا النوع يحتوي على حلقة تغذية خلفية واحدة على الأقل (مسار مغلق من التغذية الخلفية feedback loop) كما في الشكل (٣)، حيث أنه وعلى عكس الشبكة متعددة الطبقات ذات التغذية الأمامية يتم حساب الخطأ بين مخرجات الشبكة والمخرجات الحقيقية ولكن يتم تغذية ألخطأ في اتجاه عكسي (عكس اتجاه المدخلات) للشبكة لتعديل الأوزان وبالتالي يتم تخفيض الخطأ، وهذا الإجراء يتم تكراره حتى يصل الخطأ إلى أقل قيمة ممكنة.



شكل (٣)الشبكة متعددة الطبقات ذات التغذية مرتدة

(١-٧) الانحدار الحصين Robust Regression:

يؤدي وجود القيم الشاذة بالبيانات الي عدم تطابق توزيع العينة مع التوزيع الطبيعي، مما يضعف من كفاءة المقدرات التقليدية لنموذج الانحدار.

من هنا تتأتي أهمية الاستعانة بالمقدرات الحصينة لنموذج الانحدار حيث أنها أقل تأثرا بالقيم الشاذة. وهناك العديد من الطرق لايجاد المقدرات الحصينة نعرض لبعضها فيما يلي:

١-مقدرات M الحصينة:

قدم [Huber] هذه الطريقة عام ١٩٦٤ والتي يتلخص عملها في تحجيم تأثير قيم البواقي الكبير وذلك باستخدام المعادلة:

$$\phi_{e_i} = \frac{\psi_{e_i}}{e_i} \ldots\ldots\ldots\ldots\ldots\ldots (1)$$

حيث:

وفي حالة $|e_i| \to \infty$ فان قيمة ϕ_{e_i} تكون صغيرة وتقترب من الصفر وعنها يمكن استخدام أسلوب المربعات الصغري الذي يعتمد علي تصغير قيمة مجموع مربعات الخطأ الي أقل قيمة وذلك وفقا للعلاقة:

$$\hat{\beta}_M = (x' \phi x)^{-1} x' \phi x \ldots\ldots\ldots\ldots\ldots\ldots (2)$$

حيث

$$\phi = \begin{bmatrix} \phi_{e_1} & 0 \ldots\ldots & 0 \\ 0 & \phi_{e_2} \ldots\ldots & 0 \\ 0 & 0 \ldots & \phi_{e_n} \end{bmatrix}$$

٢-طريقة مربعات الوسيط الصغري:

قدم (Russecuw) هذه الطريقة عام ١٩٨٤ وذلك بغرض زيادة حصانة المقدرات في نموذج الانحدار عن طريق استبدال مجموع مربعات البواقي في طريقة المربعات الصغري العادية بوسيط مربعات البواقي، فيما يعرف بمقدرات LMS

ويشترط لسلامة التقدير في هذه الطريقة يكون عدد المتغيرات المفسرة في نموذج الانحدار أقل من نصف عدد قيم العينة أي:

$$P \leq \frac{n}{2}$$

حيث P هي المتغيرات المفسرة.

بفرض أن

$$\beta=(\beta_1,\beta_2,.....\beta_p)'$$

باستخدام أسلوب مربعات الوسيط الصغري (LMS) يمكن ايجاد b_{LMS}:

$$b_{LMS}=\min_{b} med_{i=1}^{n}(y_i-x'\beta)^2.....(3)$$

أو $b_{LMS}=\min_{b} med_{i=1}^{n} e_i^{\,2}.....(4)$

$$e_i^{\,2}=(y_i-x'\beta)^2$$

حيث:

ويسمي b_{LMS} بمقدر مربعات الوسيط الصغري، ويمثل هذا المقدار منتصف أقصر نصف في العينة الجزئية (shortest half).

ولتوضيح ذلك لنفرض أن لدينا متغير تفسيري وحيد هو x يمكن تمثيله بالمتجه

$$(x_1,x_2,.....x_n)$$

يمكن الحصول علي العينات الجزئية التالية:

$x_1, x_2, ... x_h$

$x_2, x_3, ... x_{h+1}$

$x_3, x_4, ... x_{h+2}$

$x_{n-h+1}, ... x_n$

$$h=\frac{n}{2}+1$$

x_i **هي القيمة المرتبة**

ويتم الحصول علي أصغر الفروقات كالتالي:

$$d_1=x_h-x_1$$

$$d_2=x_{h+1}-x_2$$

.

$$d_m = x_n - x_{n-h+1}$$

ثم يتم اختيار أصغر الفروقات، وتكون العينة الجزئية التي تقابله هي أقصر نصف (shortest half) بين كل العينات الجزئية الممكنة، ويكون المقدار b_{LMS} مساويا لنقطة الوسط Mid-point هذه العينة الجزئية.

٢- طريقة المربعات المشذمة:

قُدمت هذه الطريقة أيضا عن طريق (Russecuw) عام ١٩٨٤ وذلك لأنها تتميز باستقرار موضعي أفضل من طريقة مربعات الوسيط الصغرى. وتعتمد هذه الطريقة علي ترتيب مربعات الأخطاء عوضا عن ترتيب بيانات العينة. ويمكن الحصول علي مقدر b_{LTS}:

$$b_{LTS} = \min_{b} \sum_{i=1}^{h} e^{2}_{nn} \text{.....(5)}$$

حيث e^{2}_{nn} هي مربع الأخطاء المرتبة بحيث:

$$e^2_{1n} \geq e^2_{2n} \geq e^2_{3n} \geq \text{.....} \geq e^2_{nn}$$

ويمكن الحصول علي مربعات الأخطاء من العينات الجزئية السابق الاشارة اليها عن طريق حساب المتوسط لكل عينة جزئية كالتالي:

$$\overline{x}_1 = \frac{1}{h} \sum_{i=1}^{h} x_i$$

$$\overline{x}_2 = \frac{1}{h} \sum_{i=2}^{h+1} x_i$$

.

$$\overline{x}_{n-h+1} = \frac{1}{h} \sum_{n-h+1}^{n} x_i$$

بعد ذلك يتم الحصول علي مجموع المربعات لكل عينة جزئية كالتالي:

$$SQ_1 = \sum_{i=1}^{h}(x_i - \bar{x}_1)^2$$

$$SQ_2 = \sum_{i=2}^{h+1}(x_i - \bar{x}_2)^2$$

.

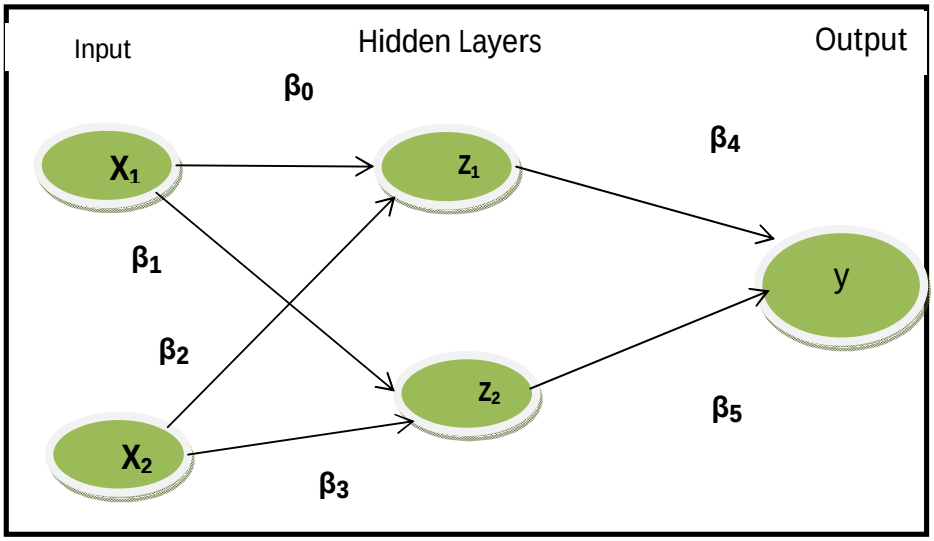
$$SQ_{n-h+1} = \sum_{n-h+1}^{n}(x_i - \bar{x}_{n-h+1})^2 \text{.....(6)}$$

ويمكن الحصول علي المربعات المشرذمة الصغري b_{LTS} هو ذلك المتوسط الذي يقابل أصغر مجموع مربعات في المعادلة رقم (٦)

(١-٨) الشبكات العصبية كنموذج انحدار:

من الممكن اعتبار الشبكات العصبية الاصطناعية كنموذج انحدار غير خطى يتكون من المتغيرات المدخلة وتتمثل في القيم الماضية yt+1,yt+2,....,yt+p أو تحويلاتها و قيم متغير خارجي xt+1,xt+2,.....,xt+h والأخطاء:

وكما في الشكل (٤) تمثل X متغيرات مستقلة داخلة في البرنامج "Input"، وتمثلZ الطبقات المخفية والتي تم فيها تدريب الشبكة "Hidden Layers"، وتمثل Y المتغير التابع والذي يقوم بإخراج البيانات بصورة جاهزة "Output".



شكل (٤) نموذج انحدار كشبكة عصبية اصطناعية

من الممكن اعتبار الشبكات العصبية الاصطناعية كنموذج انحدار:

$$Y = X\beta + \varepsilon \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (4-1)$$

حيث:

Y: متجه عمود من الدرجة $(n \times 1)$ تمثل مشاهدات عن المتغير التابع

X: مصفوفة من الدرجة $(n \times k)$ تمثل مشاهدات المتغيرات المفسرة وعددها $(k - 1)$.

β: متجه عمود يمثل معالم المجتمع وعددها k.

ε : متجه عمود من الدرجة $(n \times 1)$ يمثل الأخطاء العشوائية.

حيث يفترض أن $E(\varepsilon / X) = 0$ ، وكذلك الدالة Y غالبا ما تكون غير معلومة ويتم تقريبها بواسطة الدالة g(X,W) ، حيث W تمثل متجه المعالم وتسمى الأوزان ، والتي تكون عنصر في فراغ المعالم parameter space وتقدر من مجموعة البيانات المدربة (العينةsample) ، وكل من فراغ المعالم ومتجه المعالم يعتمد على دالة التقريب التي تم اختيارها g(X,W) وعند التنبؤ بقيمة Z مع وجود دالة التقريب g(X,W)

فانه يوجد نوعين من الخطأ ، الخطأ الأول: هو الخطأ العشوائي ε ، والخطأ الثاني: هو خطأ التقريب و الذي يمثل الفرق بين القيمة الفعلية والقيمة المقدرة ، والخطأ الثاني يكافئ الخطأ في نماذج الانحدار غير الخطية التي تحتوى على خطأ في التحديد وتوزيع الخطأ فيها يكون غير محدد، وذلك على عكس العديد من النماذج الإحصائية.

وتحديد المعالم لدالة التقريب g(X,W) تسمى ببناء الشبكة " network architecture"، وهذه الدالة عبارة عن توليفة من الدوال الخطية وغير الخطية، وعلى ذلك فإن نموذج الشبكة العصبية المقرب من الممكن أن يكون في الصيغة التالية:

$$Z = g(X, W) + \varepsilon \qquad \textbf{(4-2)}$$

حيث X تمثل مدخلات الشبكة أو طبقة المدخلات Input layer و y تمثل مخرجات الشبكة أو طبقة المخرجات output layer ، وبافتراض وجود طبقة مخفية واحدة.

أي أنه يمكن كتابة النموذج الخطي كما يلي:

$$Y = Z.W_2 + \varepsilon_2 \quad \text{.....} \quad (4-3)$$

حيث: $W_2 = \begin{bmatrix} \beta_4 \\ \beta_5 \end{bmatrix}$

$$Z = X.W_1 + \varepsilon_1 \quad \text{.....} \quad (4-4)$$

$W_1 = [\beta_0 \quad \beta_1 \quad \beta_2 \quad \beta_3]$, $X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

حيث:

W, β , W_1 , W_2: معالم النموذج (أوزان الشبكة).

وبالتعويض عن Z في الدالة Y نحصل على:

$$Y = (X W_1 + \varepsilon_2) W_2 + \varepsilon_1$$

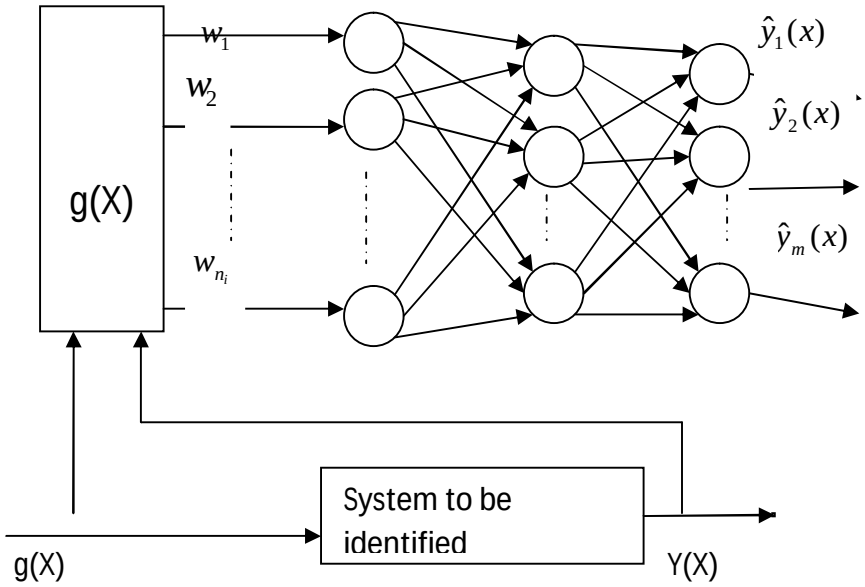
$$Y = X W_1 W_2 + \varepsilon_2 W_2 + \varepsilon_1 \quad \text{.....} \quad (4-5)$$

وبوضع: $\varepsilon_2 W_2 + \varepsilon_1 = \varepsilon\beta = W_2 W_1$,

نستنتج أن:

$$Y = X\beta + \varepsilon_1 \quad \ldots\ldots \quad (4-1) \quad)$$

وبالتالي نكون قد حصلنا على النموذج العام للانحدار المتعدد.ومن الممكن إضافة أكثر من طبقة مخفية وسنحصل على نموذج مشابه للنموذج (٤-١) السابق والهيكل البنائي للشبكة المستخدم في هذه الدراسة هو الشبكة متعددة الطبقات ذات التغذية الأمامية Multi-layer FeedForward Neural Network (MFNN) ، حيث تستطيع الشبكة متعددة الطبقات ذات التغذية الأمامية حل العديد من المشاكل المعقدة التي لا تستطيع الشبكة ذات الطبقة الواحدة حلها ، كما أن هيكل هذه الشبكة من أكثر هياكل الشبكات العصبية استخداما (Shiraz Amjad,2003).



شكل (٥) الهيكل العام للشبكة العصبية متعددة الطبقات ذات التغذية الأمامية.

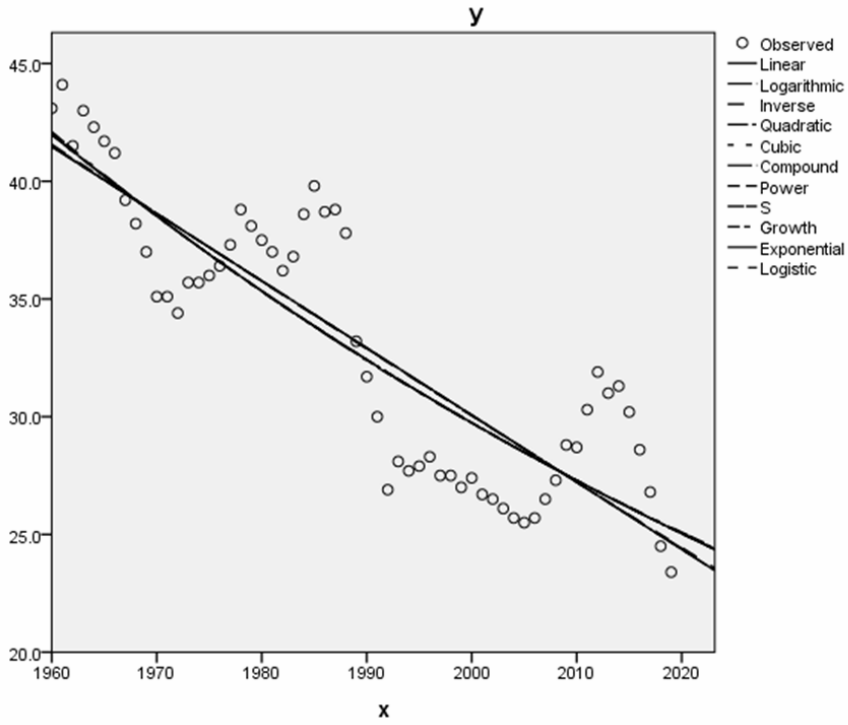
(١-٩) الدراسة تطبيقية

عمدت الكثير من الأبحاث الي التنبؤ بمعدلات المواليد لما له من أهمية في التحكم بهذه الظاهرة، والتخطيط لمجابهة تبعاتها، تعليميا وصحيا و علي مستوي الرعاية الاجتماعية، فضلا عن حسن استغلال الهبة الديموجرافية المتوقعة.

وبالنسبة لهذا البحث الذي يدرس معدلات المواليد للسكان في مصر فقد تم اختيار عينة كبيرة حجمها ٦٠ مفردة، حيث تمتد السلسة الزمنية منذ عام ١٩٦٠ حتى عام ٢٠١٩. وقد تم اختيار عام ١٩٦٠ كبداية للسلسلة لأنه بداية فترة الستينات التي شهدت بداية انخفاض معدلات المواليد في مصر.

(١-٩-١) النماذج التقليدية:

استخدم الباحثان نماذج الاندار التقليدية سواء الخطية أو اللوغاريتمية أو الأسية والتي قدرت معالمها بطريقة المربعات الصغري، وجميعها أعطت معاملات تحديد صغيرة حيث تراوحت بين (٠.٧٢٩ – ٠.٧٣٩) مما يدل علي عدم جودة توفيق البيانات لجميع هذه النماذج. الشكل التالي يوضح ذلك:



يتضح من الشكل السابق أن بيانات المواليد تعاني من وجود قيم شاذة تتناثر بعيدا عن الخطوط الممثلة لنماذج الانحدار التقليدية، كما اتضح من اختبار طبيعية البيانات عدم تبعية البيانات للتوزيع الطبيعي كما هو موضح.

(١-٩-٢) اختبار طبيعية البيانات:

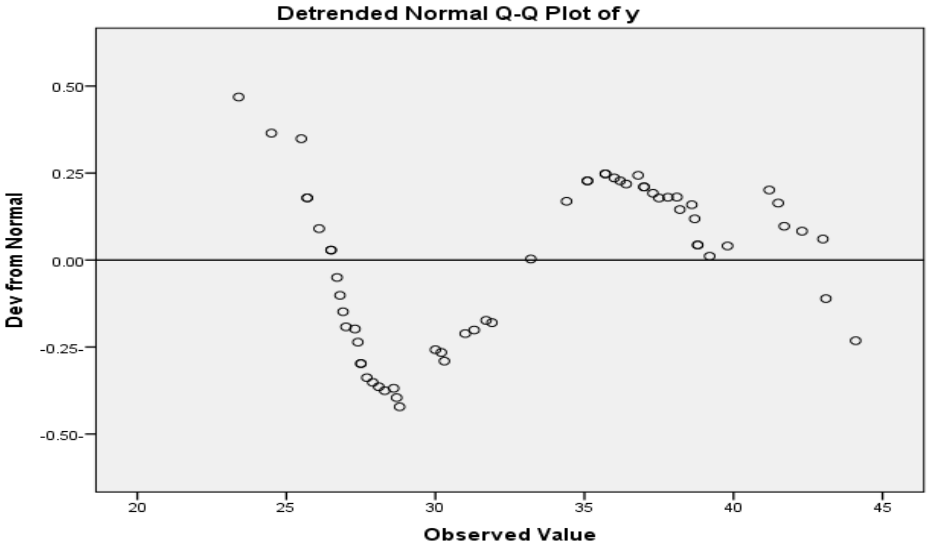
نقوم باستخدام اختبار كولمجروف وسيمرنوف وذلك لأن العينة كبيرة الحجم (٥٠<) وذلك لاختبار الفرض العدمي القائل بان البيانات تتبع توزيعا طبيعيا في مقابل الفرض البديل القائل بأن البيانات لا تتبع توزيعا طبيعيا.

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Y	.152	60	.001	.932	60	.002

a. Lilliefors Significance Correction

يظهر جدول النتائج أن قيمة (Sig) في اختبار (Kolmogorov-Smirnova) هي (٠.٠٠١) وهي قيمة تقل عن مستوي المعنوية (٠.٠٥) بالتالي نرفض الفرض العدمي القائل بأن البيانات تتبع التوزيع الطبيعي. بالتالي فان البيانات لا تتبع التوزيع الطبيعي. لذا لجأ الباحث الي استخدام نموذج الانحدار الحصين



(١-٩-٣) نموذج الانحدار الحصين:

استخدم الباحث برنامج R في تقدير نموذج الانحدار غير الخطي بطريقة الامكان الأعظم وغير الخطي الحصين بطريقة algorithm وكانت النتائج كالتالي:

Parameters	Estimate	Std. Error	t value	Pr(>|t|)
Nonlinear regression				
A	26.34854	7.62174	3.457	0.00104 **
B	0.01867	0.01016	1.838	.0.07130
C	17.40874	8.46436	2.057	0.04430 *
AIC(m)= 305.95		**BIC**(m)= 314.33		
Nonlinear robust regression				
A	25.32908	8.67304	2.920	0.0050 **
B	0.02017	0.01372	1.470	0.1470
C	18.50809	9.77990	1.892	.0.0635

AIC(m)= 303.24		**BIC**(m)= 308.13		
Simultaneous Tests for General Linear Hypotheses				
A	25.32898	8.67289	2.920	0.00398 **
B	0.02017	0.01372	1.470	0.17533
C	18.50822	9.77974	1.893	.0.07543
Signif. codes: '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				

يتبين من الجدول السابق أن نموذج الانحدار غير الخطي أعطي دلالة معنوية للمقدر A عند مستوي معنوية ٠.٠١ وللمقدر B عند مستوي معنوية ٠.١ وللمقدر Cعند مستوي معنوية ٠.٠٥، بينما النموذج الحصين أعطي دلالة معنوية للمقدر A عند مستوي معنوية ٠.٠١ وللمقدر Cعند مستوي معنوية ٠.١ ولم يعطي دلالة معنوية للمقدر B وبذلك يكون نموذج الانحدار الحصين كالتالي:

$$Y = 25.329e^{-X} + 18.508$$

ويكون نموذج الانحدار غير الخطي كالتالي:

$$Y = 26.349e^{-0.0187X} + 17.409$$

كما تبين من الجدول السابق أن نموذج الأنحدار الحصين أفضل من نموذج الانحدار غير الخطي الحصين وذلك باستخدام معياري AIC للمعلومات و BIC، إلا أن متوسطات الخطأ المعياري لمقدرات نموذج الانحدار غير الخطي أقل من مثيلاتها لدي نموذج الانحدار الحصين.

(١٠-١) طريقة الشبكات العصبية الاصطناعية:

(١-١٠-١) توصيف النموذج بطريقة الشبكة العصبية:

يتم توصيف النموذج بطريقة الشبكة العصبية عن طريق:

١- دراسة بيانات المدخلات والمخرجات

الخطوة الأولى لتحليل البيانات هي تهيئة هذه البيانات لكي تصبح صالحة لإجراء العمليات الحسابية عليها ولكي نحصل منها على تنبؤات دقيقة. وتتم هذه التهيئة على مرحلتين: مرحلة الفصل (Separation) و مرحلة التحويل أو المعايرة (Scaling).

مرحلة الفصل: وهى مرحلة خلق مجموعة البيانات الخاصة بمرحلة التدريب ومرحلة الاختبار، وهى مرحلة مهمة لتحديد فعالية النماذج الإحصائية و نماذج الشبكات العصبية الاصطناعية من خلال تزويد كل من النموذجين ببيانات مدخلة ، ومن ثم مراقبة ما إذا كانت مخرجات كل نموذج تعطى النتائج المتوقعة ، وفى سبيل تحقق ذلك يتم تقسيم البيانات المتاحة إلى مجموعتين:

١-مجموعة التدريب: وهذه المجموعة تحتوى على n-3 مشاهدة تستخدم للحصول على النموذج المناسب.

٢-مجموعة الاختبار: وهذه المجموعة تحتوى على الثلاثة مشاهدات الأخيرة. ويجب التنبيه على إن المجموعة التي تستخدم في التدريب لا تستخدم في الاختبار.

مرحلة التحويل: تستخدم في غالبية نماذج الشبكات العصبية دوال رياضية لتحويل البيانات المخرجة من هذه الشبكات ،حيث تنحصر القيم في الفترة ٠,١ ، وذلك بغرض استخدام بعض الدوال المنطقية (Logistic Function) عليها كدوال تفعيلية

٢- هيكل بناء الشبكة network architecture:

من الممكن اعتبار الشبكات العصبية الاصطناعية كنموذج انحدار غير خطى

$$Z = v(X) + \varepsilon \qquad \textbf{(1)}$$

حيث يفترض أن $E(\varepsilon / X) = 0$ **، وكذلك الدالة** $v(x) = E(Z / X)$ **غالبا ما تكون غير معلومة ويتم تقريبها بواسطة الدالة g(X,W) ، حيث W تمثل متجه المعالم وتسمى الأوزان ، والتي تكون عنصر في فراغ المعالم parameter spaceوتقدر من مجموعة البيانات المدربة (العينةsample) ، وكل من فراغ المعالم ومتجه المعالم يعتمد على دالة التقريب التي تم اختيارها g(X,W)، وعلى ذلك فان نموذج الشبكة العصبية المقرب من الممكن أن يكون في الصيغة التالية:**

$$Z_{t+1} = g(X, W) + \varepsilon_{t+1}$$

والهيكل البنائي للشبكة المستخدم في هذه الدراسة هو الشبكة متعددة الطبقات ذات التغذية الأمامية Multi-layer FeedForward Neural Network (MFNN))، وذلك لأن الشبكة متعددة الطبقات ذات التغذية الأمامية تستطيع حل العديد من المشاكل المعقدة التي لا تستطيع الشبكة ذات الطبقة الواحدة حلها. كما أن هيكل هذه الشبكة من أكثر هياكل الشبكات العصبية استخداما في النواحي التطبيقية.

(١-١٠-٢) تقدير معالم النموذج Estimation of Model parameters

يقصد به تقدير أوزان الشبكة (معالم النموذج) باستخدام أسلوب الشبكات العصبية الاصطناعية ، و لكي يتم ذلك يلزم جعل متوسط مربعات الخطأ MSE)) اقل ما يمكن. و من الممكن الحصول على تقدير أوزان الشبكة (معالم النموذج) عن طريق تدريب الشبكة باستخدام مجموعة بيانات التدريب ، حيث تمثل هذه الأوزان المعلومات الأولية التي ستتعلم بها الشبكة. لذا لا بد من تحديث الأوزان خلال مرحلة التدريب ، ومن أجل هذا التحديث تستخدم عدة خوارزميات مختلفة (داخل حزم الماتلاب) حسب نوع الشبكة ، من أهم هذه الخوارزميات خوارزمية الانتشار العكسي Back Propagation Algorithm التي تستخدم في تدريب الشبكات العصبية متعددة الطبقات ذات التغذية الأمامية وغير الخطية، وتتم عملية تدريب الشبكة داخل برنامج MATLAB بالتسلسل التالي:

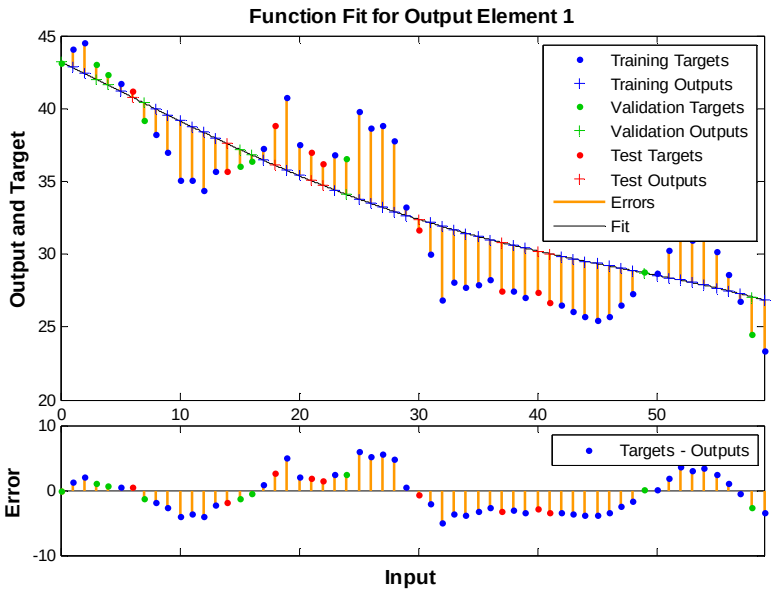
(١) إنشاء الشبكة network creation:

إن أول خطوة في تدريب الشبكة هو إنشاء الشبكة، وبما أننا نريد إنشاء شبكة أمامية بتغذية عكسية للخطأ ، فسوف نستخدم الأمر الآتي:

('net = fitnet(2,'trainlm

(٢) التدريب training:

وخلال التدريب تتغير هذه الأوزان والانحياز بشكل تكراري إلى الوصول إلى القيمة الصغرى لتابع الكلفة أو ما يسمى تابع الأداء performance function. إن تابع الأداء الافتراضي لشبكات التغذية الأمامية هو متوسط مربع الخطأ (MSE).



وفى أغلب تطبيقات الشبكات العصبية يتم تحديد تلك الأوزان من خلال تصغير مجموع مربعات أخطاء الانحدار (البواقي) (Sum of Squared Errors (SSE.

وسوف يتم استخدام أسلوب التغذية الخلفية Back propagation algorithm

من أكثر الأساليب التدريبية استخداما في تدريب الشبكة العصبية ، ولقد تم تقديم هذا الأسلوب لأول مرة من خلال (Paul Werbos,1974) ، ثم تم تطويره بواسطة (Rumelhart,Hinton,Williams,1986) ، من خلال بناء أول شبكة عصبية لها أكثر من طبقة خفية. والهدف الأساسي من استخدام هذا الأسلوب في التدريب هو حساب الخطأ بين المخرجات المستهدفة والمخرجات الفعلية في طبقة المخرجات ثم إعادته مرة أخرى لكل طبقة خفية من أجل تعديل أوزان الاتصال حتى يتم الوصول إلى طبقة المدخلات، ويتم هذا التعديل في الأوزان إلى أن تحقق الشبكة أقل مجموع مربع أخطاء عند التدريب (Dhar and Stein,1996).

(١-١٠-٣) التحقق من صلاحية النموذج Checking of the Model

تعد هذه المرحلة من المراحل الأساسية في التحليل باستخدام الشبكات العصبية الاصطناعية ، حيث أنه على أساسها يتم تحديد مدى قبول النموذج الذي تم توصيفه وتقدير معالمه، ووفقاً لهذه المرحلة يحدد إما بالاستمرار في عملية التحليل وتحقيق ما هو مستهدف من نموذج التحليل أو العودة إلى نقطة البداية من تحديد وتقدير ثم اختبار ، ويتم اختيار النموذج الذي يحقق أقل متوسط مربعات أخطاء MSE حيث متوسط مربعات الأخطاء يحسب كما يلي:

$$MSE = \frac{\sum(y_i - \hat{y})^2}{n} \qquad (2)$$

حيث:

y: المشاهدات المستهدفة **Target observation**.

$\hat{y}$: القيمة التنبؤية **Predicted value**.

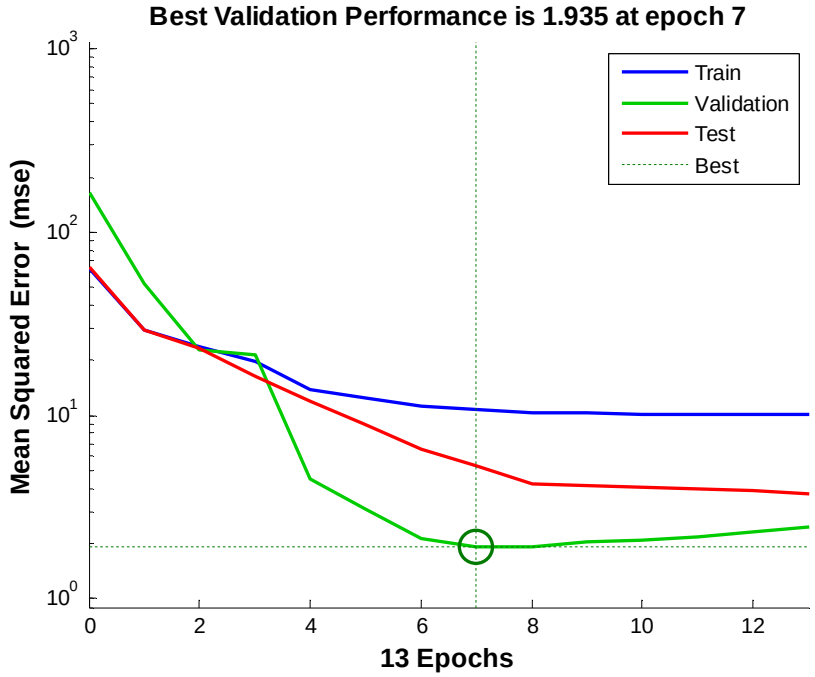
n: عدد المشاهدات. **Number of observations**

كذالك مقياس متوسط مجموع الانحرافات المطلقة **Mean Absolute Deviation (MAD)**:

$$Mean\ absolute\ error\ (MAE) = \frac{\sum_{i=1}^{n} | y_i - \hat{y}_i |}{n} \qquad (3)$$

ولقد تم حساب مقياس متوسط مربعات الخطا MSE ومتوسط الانحرافات المطلقة لكل من الاسلوبين (الانحدار الحصين – الشبكات العصبية الاصطناعية) باستخدام برنامج MATLAB وبرنامج R (انظر الملحق)

بالإضافة إلى معايير تشمل عدد مرات إجراء تعديل الأوزان (معالم النموذج) وتسمى Epoch والتي يمكن التحكم بها في النموذج (وكما في الشكل التالى تم اجراء Epoch 13 للوصول إلى افضل اداء للشبكة لتحقيق اقل اخطاء ممكنة والذى تحقق عند Epoch 7).



(١-١٠-٤) استخدام النموذج في التنبؤ و/أو التحكم:

بعد التأكد من صلاحية النموذج، يتم استخدامه في التنبؤ بالمشاهدات المجهولة للظاهرة. وتقاس كفاءة النماذج الإحصائية بمدى قدرتها على تمثيل الواقع وقدرتها على التنبؤ بالقيم المجهولة ، وتعد النماذج الأقل خطأ في ذلك هي النماذج الأفضل. أما بالنسبة لمقارنة النماذج الجيدة فيتم باستخدام خطأ التنبؤ الكلى المقدر باستخدام تلك النماذج.

وسوف يتم استخدام برنامج MATLAB الذي يعد أفضل البرامج في مجال تدريب واختبار الشبكات كما يتضمن أكواد مناسبة لإجراء التحليل التقليدي لنماذج الانحدار المتعدد ذلك لأن برنامج MATLAB يتضمن ميزة إمكانية استخدامه لإجراء تحليل نماذج الانحدار المتعدد بالطريقة التقليدية وبالطريقة المقترحة باستخدام الشبكات العصبية. لتفاصيل أكثر عن برنامج MATLAB يرجى الدخول على www.Matlab.com.

ويتم دراسة هذه المرحلة كما يلي:

عند التنبؤ بقيم الانحدار المتعدد باستخدام الشبكات العصبية الاصطناعية بالاستعانة ببرنامج MATLAB تم استخدام الدالة التالية:

(Y = sim(net ,x) ..... (4

حيث:

Y: تمثل القيم التنبؤية للثلاثة قيم المجهولة باستخدام أسلوب الشبكات العصبية الاصطناعية.

net: تمثل الشبكة المدربة.

X: تمثل البيانات التي تم استخدامها لتدريب الشبكة والتي حجمها هو n-3.

(١-١١) المقارنة بين نتائج الانحدار الحصين والذكاء الاصطناعي:

بعد حساب القيم التنبؤية لقيم معدلات المواليد بجانب المشاهدة الثلاثة المستقبلية عن طريق دوال MATLAB المشار إليها سابقا للشبكات العصبية وكذلك برنامج Rلاسلوب الانحدار الحصين ، تم حساب دقة هذه التنبؤات وذلك بواسطة مقاييس إحصائية تعتمد على الفرق بين القيمة التنبؤية و القيمة الفعلية (البواقي) وهى:

ومن الجدير بالذكر أن المقاييس التى تم استخدامها في هذه الدراسة هي (MSE,MAD) ، وذلك لشيوع استخدامها في تحليل الانحدار.

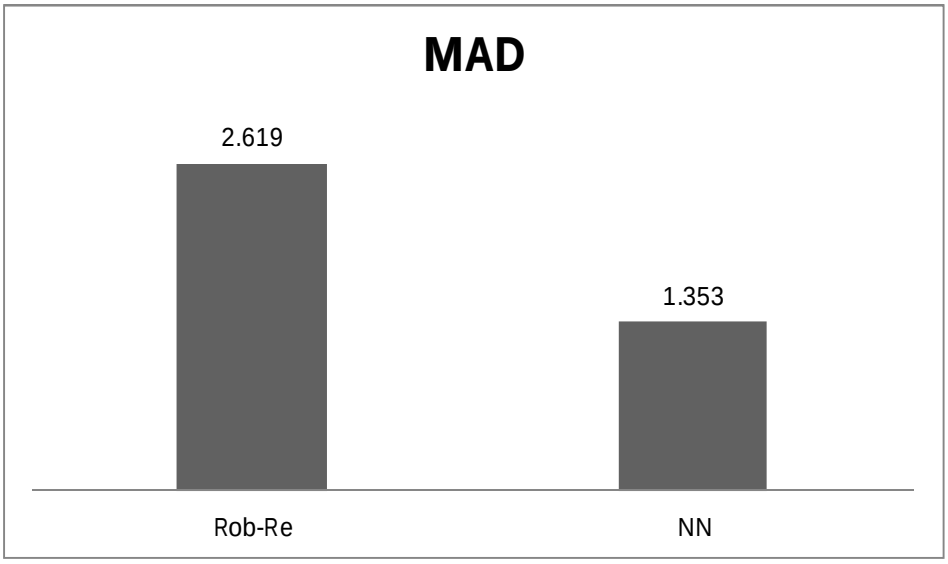
ولقد تم الحصول على نتائج (MSE,MAD) لتنبؤات الشبكات العصبية NN)) وتنبؤات الانحدار الحصين Rob-Re)) كانت كما في جدول (١) التالى:

Year	Y	Y(predict)		MAD		MSE	
		NN	Rob-Re	NN	Rob-Re	NN	Rob-Re
1960	43.1	43.35211	41.61	0.252112	1.49	0.06356	2.2201
1961	44.1	43.34385	41.32	0.75615	2.78	0.571763	7.7284
1962	41.5	43.31596	41.03	1.184036	0.47	1.401941	0.2209
1963	43	43.22319	40.74	0.22319	2.26	0.049814	5.1076
1964	42.3	42.92872	40.45	0.628718	1.85	0.395286	3.4225
1965	41.7	42.118	40.15	0.418004	1.55	0.174727	2.4025
1966	41.2	40.54724	39.86	0.652758	1.34	0.426094	1.7956
1967	39.2	38.87102	39.57	0.328983	0.37	0.10823	0.1369
1968	38.2	37.91464	39.28	0.285357	1.08	0.081429	1.1664

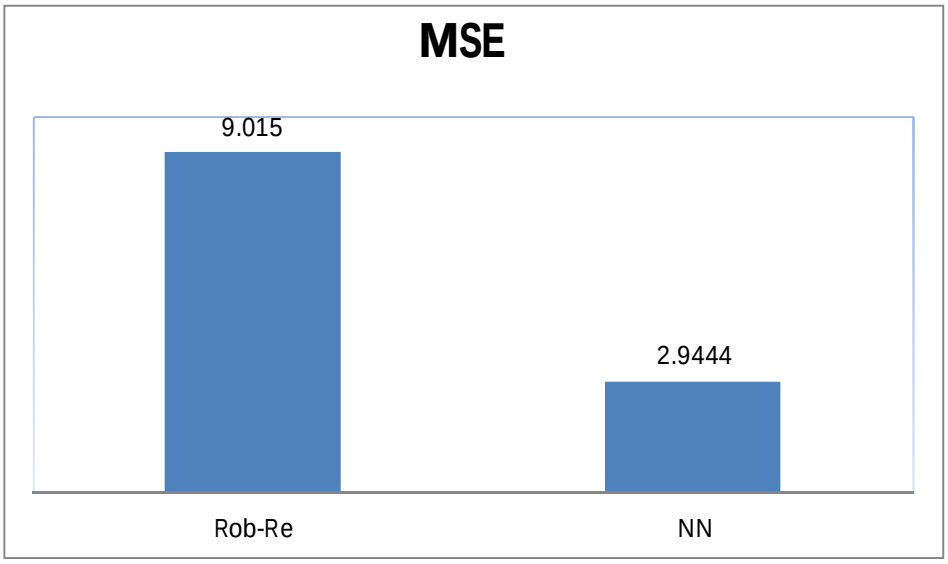
1969	37	37.55004	38.99	0.550039	1.99	0.302543	3.9601
1970	35.1	37.43317	38.70	2.333165	3.6	5.443661	12.96
1971	35.1	37.39785	38.41	2.297845	3.31	5.280092	10.9561
1972	34.4	37.38736	38.12	2.987363	3.72	8.924339	13.8384
1973	35.7	37.38427	37.83	1.684269	2.13	2.836763	4.5369
1974	35.7	37.38336	37.54	1.683358	1.84	2.833693	3.3856
1975	36	37.38309	37.25	1.383089	1.25	1.912935	1.5625
1976	36.4	37.38301	36.96	0.98301	0.56	0.966308	0.3136
1977	37.3	37.38299	36.67	0.082987	0.63	0.006887	0.3969
1978	38.8	37.38298	36.38	1.41702	2.42	2.007947	5.8564
1979	40.8	37.38298	36.09	3.417023	4.71	11.67605	22.1841
1980	37.5	37.38297	35.80	0.117027	1.7	0.013695	2.89
1981	37	37.38296	35.51	0.382957	1.49	0.146656	2.2201
1982	36.2	37.38289	35.22	1.182887	0.98	1.399223	0.9604
1983	36.8	37.38257	34.93	0.582567	1.87	0.339385	3.4969
1984	38.6	37.3811	34.63	0.781101	3.97	0.610119	15.7609
1985	39.8	37.3744	34.34	2.425604	5.46	5.883555	29.8116
1986	38.7	37.34382	34.05	1.356181	4.65	1.839227	21.6225
1987	38.8	37.20635	33.76	1.59365	5.04	2.53972	25.4016
1988	37.8	36.62562	33.47	1.17438	4.33	1.379169	18.7489
1989	33.2	34.69605	33.18	1.496049	0.02	2.238163	0.0004
1990	31.7	31.32484	32.89	0.37516	1.19	0.140745	1.4161
1991	30	29.03796	32.60	0.962044	2.6	0.925528	6.76
1992	26.9	28.28808	32.31	1.388081	5.41	1.926768	29.2681
1993	28.1	28.106	32.02	0.006004	3.92	3.60E-05	15.3664
1994	27.7	28.06526	31.73	0.365261	4.03	0.133415	16.2409
1995	27.9	28.05631	31.44	0.156314	3.54	0.024434	12.5316
1996	28.3	28.05436	31.15	0.245642	2.85	0.06034	8.1225
1997	27.5	28.05393	30.86	0.55393	3.36	0.306839	11.2896
1998	27.5	28.05384	30.57	0.553837	3.07	0.306735	9.4249
1999	27	28.05382	30.28	1.053816	3.28	1.110529	10.7584
2000	27.4	28.05381	29.99	0.653812	2.59	0.42747	6.7081
2001	26.7	28.05381	29.70	1.353811	3	1.832804	9
2002	26.5	28.05381	29.41	1.553811	2.91	2.414328	8.4681
2003	26.1	28.05381	29.11	1.953811	3.01	3.817376	9.0601
2004	25.7	28.05381	28.82	2.353811	3.12	5.540425	9.7344
2005	25.5	28.05381	28.53	2.553811	3.03	6.521949	9.1809
2006	25.7	28.05381	28.24	2.353811	2.54	5.540425	6.4516
2007	26.5	28.05381	27.95	1.553811	1.45	2.414328	2.1025
2008	27.3	28.05381	27.66	0.753811	0.36	0.568231	0.1296
2009	28.8	28.05381	27.37	0.746189	1.43	0.556798	2.0449
2010	28.7	28.05381	27.08	0.646189	1.62	0.417561	2.6244
2011	30.3	28.05381	26.79	2.246189	3.51	5.045366	12.3201
2012	31.9	28.05381	26.50	3.846189	5.4	14.79317	29.16
2013	31	28.05381	26.21	2.946189	4.79	8.680031	22.9441
2014	31.3	28.05381	25.92	3.246189	5.38	10.53774	28.9444
2015	30.2	28.05381	25.63	2.146189	4.57	4.606128	20.8849
2016	28.6	28.05381	25.34	0.546189	3.26	0.298323	10.6276
2017	26.8	28.05381	25.05	1.253811	1.75	1.572041	3.0625
2018	24.5	28.05381	24.76	3.553811	0.26	12.62957	0.0676
2019	23.4	28.05381	24.47	4.653811	1.07	21.65795	1.1449
Average				1.353	2.619	2.9444	9.015

جدول (١) نتائج (MSE,MAD) لتنبؤات الشبكات العصبية (NN) وتنبؤات الانحدار الحصين (Rob-Re)

من خلال الجدول السابق يتضح لنا أن جميع قيم MAD,MSE للشبكات العصبية الاصطناعية كانت اقل من أسلوب الانحدار الحصين كذلك كان المتوسط العام لمربعات الأخطاء للشبكات MSE(NN)=2.944 بينما كان MEE(Rob-Re)=9.015 للانحدار الحصين مما يعنى أفضلية أسلوب الشبكات العصبية الاصطناعية عن الانحدار الحصين للتنبؤ بمعدلات المواليد.



شكل(٦)يوضح قيم متوسط الأخطاء المطلق للانحدار الحصين والشبكات العصبية



شكل(٧) يوضح قيم متوسط مربعات الأخطاء للانحدار الحصين والشبكات العصبية

(١-١٢) النتائج والتوصيات

أولاً: النتائج

١- توجد علاقة غير خطية بين معدلات المواليد والزمن حيث تتناقص معدلات المواليد بمرور الزمن بصورة غير خطية، إلا أن النموذج غير الخطي الذي تم تقدير معالمه باستخدام طريقة المربعات الصغري لا يتمتع بالكفاءة في تمثيل بيانات المواليد التي لا تتبع توزيعا طبيعيا وتحوي قيما شاذة.

٢- استخدام الاساليب الحصينة في تقدير معالم النموذج الانحدار غير الخطي ادي الي تحسين مقدرات النموذج ومن ثم جودة تمثيله لبيانات المواليد في مصر.

٣- هناك أفضلية لاستخدام الذكاء الاصطناعي والشبكات العصبية عن النماذج الإحصائية التقليدية سواء تلك التي استخدمت طريقة المربعات الصغري أو الطرق الحصينة في تقدير معالمها، وذلك في التنبؤ بمعدلات المواليد في مصر.

ثانياً: التوصيات

١- يوصي الباحثان باستخدام النماذج غير الخطية في الإسقاطات السكانية.

٢- يوصي الباحثان باستخدام الطرق الحصينة في التعامل مع البيانات التي لا تتبع توزيعا طبيعيا.

٣- يوصي الباحثان باستخدام الشبكات العصبية الي جوار النماذج الإحصائية التقليدية في الدراسات السكانية.

أولا:المراجع العربية

(١) ايمان بكري (٢٠١٤). التنبؤ بعدد المواليد بمحلية عطبرة باستخدام السلاسل الزمنية خلال الفترة (٢٠٠٣-٢٠١٣)، جامعة وادي النيل، السودان.

(٢) جعفر محمد حاجى ، محمد عبد الهادي المحميد ،(١٩٩٩م). "الشبكات العصبية: التنبؤ بأسعار صرف الدينار الكويتي مقابل الدولار الأمريكي " ، المجلة العربية للعلوم الإدارية ، مجلد ٦ ، عدد ١ ، يناير (١٩٩٩) ، ص ١٧- ٣٥.

(٣) ميشيل نجينفيتسكى ،(٢٠٠٤م). " الذكاء الصناعى دليل النظم الذكية " تعريب سرور على إبراهيم سرور ، دار المريخ للنشر ، الرياض ، المملكة العربية السعودية، ص ٢٥٢.

ثانيا:المراجع الأجنبية

1) Arminger, G. and Enache, D. (1996)
, "Statistical Models and Artificial Neural Networks". In: Bock, H.H. and Polasek, W. (Eds.): *Data Analysis and Information Systems*, Vol. 7, Springer Verlag, Heidelberg, 243-260.

2) Al-Shawadfi , Gamal A.(1994).
"Bayesian Inference of ARMAX Models" , Scientific Magazine, Faculty of Commerce, Al-Azhar University ,Cairo, Egypt, Vol. 20 July 1994.

3) Al-Shawadfi , Gamal A.(1996).
"Bayesian Estimation for the Parameters of the Seasonal ARMAX Models", Scientific Magazine, Faculty of

Commerce, Ain Shams University Cairo, Egypt, Vol. 1 July 1996,PP. 139-151.

4) Al-Shawadfi , Gamal A.(1997).
"The Derivation of Bayesian Predictive Density for the Seasonal ARMAX Model", Scientific Magazine, Faculty of Commerce Tanta University, Vol.1 April 1997 ,Tanta, Egypt.

5) Al-Shawadfi , Gamal A.(2003)
" A comparison between neural network and Box-Jenkins Forecasting Techniques With Application to Real data " King Saud University ,King Fahd National Library Cataloging-in-Publication Data.

6) Box , G.P. and Jenkins, G., M (1976).
" Time Series Analysis Forecasting And Control" , Holden – day , San Francisco.

7) Box, G. E. P., Jenkins, G. M., and Reinsel, G. C. (1994).
" Time Series Analysis, Forecasting and Control", (3rd ed.), Englewood Cliffs, NJ: Prentice-Hal l, p (151).

8) Bosson - Amedenu Senyefia, Joseph Otoo (2019)
" Time Series Analysis and Forecasting of Caesarian Section Births in Ghana". Biomedical Statistics and Informatics ·

9) Cybenko, g.(1989).
"Approximation by superposition of sigmoidal functions ", Mathematics of control , Signals and Systems, 2, 303-314.

10) Demuth H. & Beale M. (1998).
" Neural Networks Toolbox:Users Guide, Version 3.0",The Math Works inc Natick, MA.

11) arvey , A. C. and Phillips ,G.A.(1977).
" A comparison of Estimators in the ARMA (1,1)model" ,University of Kent(Unpublished Paper)
.

12) Hijazi, D. (2007),
"Predicting the stock market method of using artificial neural cells." , M.A. Thesis Department of Statistics, Faculty of Economics and Political Science, Cairo University, Cairo, Egypt.

13) John McDonald (1979)
"ATIME SERIES APPROACH TO FORECASTING AUSTRALIAN TOTAL LIVE-BIRTHS". DEMOGRAPHY, Volume16, Number 4

14) John McDonald (1981)
"Modeling Demographic Relationships: An Analysis of Forecast Functions for Australian Births". Journal of the American Statistical Association, Vol. 76, No. 376.

15) Joop de Beer (1991)
"From birth expectations to birth forecasts: A partial‐adjustment approach", Mathematical Population Studies: An International Journal of Mathematical Demography.

16) José Leopoldo Ferreira Antunes (1998)
"Grow and multiply": social development, birth rates and demographic transition in the Municipality of São Paulo, Brazil, time-series for 1901-94". Rev. Bras. Epidemiol. Vol. 1, Nº 1.

17) Jorge Miguel Bravo and Others (2013)
"On the use of Seasonal Forecasting Methods to model birth and deaths data". Eurostat/UNECE Work Session on Demographic Projections, Rome, Italy.

18) Jorge M. Bravo, Edviges Coelho (2019)

" Modelling Monthly Births and Deaths Using Seasonal Forecasting Methods as an Input for Population Estimates".
https://www.researchgate.net/publication/333998261

19) Kenneth C. Land, David Cantor (1983)
"ARIMA MODELS OF SEASONAL VARIATION IN U. S. BIRTH AND DEATH RATES", DEMOGRAPHY, Volume20, Number 4

20) Kenneth F. Wallis (1981)
" Modeling Demographic Relationships: An Analysis of Forecast Functions for Australian Births: Comment". American Statistical Association, Journal of the American Statistical Association, Vol. 76, No. 376.

21) Lawrence R. Carter(1996)
"Forecasting U.S. Mortality: A Comparison of Box-Jenkins ARIMA and Structural Time Series Models", Wiley on behalf of the Midwest Sociological Society.

22) Peter Congdon (1980)
"Forecasting Births in Greater London: An Application of the Easterlin Hypothesis". Taylor & Francis, Ltd.

23) Robert McNown and Andrei Rogers (1989)
"Forecasting Mortality: A Parameterized Time Series Approach", Demography, Vol. 26, No.4.

24) Shaarawy S. and Ismail M. A., (1987)
" Bayesian inference for seasonal ARMA models", Egypt. Statist. J. 31, 323-336.

25) Turban, E., Aronson J. and Liang T. (2005).
"Decision Support Systems and Intelligent Systems ". 7th edition Prentice Hall.

26) Zurada ,J. M. (1992).
"Artificial Neural Systems".1st ed. St. Paul, MN., West Publishing Company.

ملحق (A)

برنامج لتقديرالمعالم لنموذج الانحدار الحصين والتنبؤ بمعدلات المواليد باستخدام حزمR.

```
rm(list = ls())
rat=c(43.1,44.1,41.5,43,42.3,41.7,41.2,39.2,38.2,37,35.1,35.1,34.4,35.7,35.7,36,36.4,
37.3,38.8,40.8,37.5,37,36.2,36.8,38.6,39.8,38.7,38.8,37.8,33.2,31.7,30,26.9,28.1,27.7,
27.9,28.3,27.5,27.5,27,27.4,26.7,26.5,26.1,25.7,25.5,25.7,26.5,27.3,28.8,28.7,30.3,31.
9,31,31.3,30.2,28.6,26.8,24.5,23.4)
dd=c(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,
30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50,51,52,53,54,55,56,5
7,58,59,60)
boxplot(rat(
#DV=c(1,4,3,6,5,8,1,4,3,6,5,12,1,2,3,4,5,6,1,2,3,4,5,6(
# IV1 = factor(c(rep("A",6), rep("B",6), rep("C",6), rep("D",6(((
# IV2 = factor(rep(c("M","N"),12((
 library(robustbase(
 model = lmrob(rat ~ dd(
 df <- data.frame(dd,r)
 m <- nls(rat ~ I(a*exp(-b*dd)+c), data=df, start=list(a=max(rat), b=1, c=10),
trace=T(
 summary(m(
 AIC(m(
 BIC(m(
 mod=nlrob(rat~I(a*exp(-b*dd)+c),data=df,start =list(a=26, b=0.02, c=17((
 mod
 summary(mod(
 fitted(m)
 plot(fitted(m))
 residuals(mod)
 plot(residuals(mod))
AIC(mod)
BIC(mod)
## robust
mr <- nlrob(rat ~ I(a*exp(-b*dd)+c), data=df, start=list(a=max(rat), b=1, c=10),
trace=T)
mr
summary(mr)
coef(summary(mr))
residuals(mr)
plot(residuals(mr))
predict(mr)
plot(predict(mr))
AIC(mr)
```

```
BIC(mr)
library(multcomp)
mcr = glht(mr)
mcrs = summary(mcr, test=adjusted("single-step"))
mcrs
## robust
mrc <- nlrob.control(rat ~ I(a*exp(-b*dd)+c), data=df, start=list(a=max(rat), b=1,
c=10), trace=T)
summary(mrc)
summary(mod)
 summary(model)
 ### Effect of IV1

 model.2 = lmrob(DV ~ IV2)

 anova(model, model.2)

 ### Effect of IV2

 model.3 = lmrob(DV ~ IV1)

 anova(model, model.3)
  library(car)
 Anova(model)
 library(multcomp)
 mc = glht(model)
 mcs = summary(mc, test=adjusted("single-step"))
 mcs
 mc = glht(model,mcp(IV1 = "Tukey"))
 mcs = summary(mc, test=adjusted("single-step"))
 mcs
 plot(dd,rat,xlim=c(0,60),ylim=c(0,50))
 fit1<-lmrob(rat ~ dd)
 abline(fit1$coef,lty=1)
 opar <- par(mfrow = c(2,2), oma = c(0, 0, 1.1, 0))
 plot(model, las = 1)
 par(opar)
 d1 <- cooks.distance(model)
 r <- stdres(model)
 #exponintial regression
 df <- data.frame(x, y)
 m <- nls(y ~ I(a*exp(-b*x)+c), data=df, start=list(a=max(y), b=1, c=10), trace=T)
 y_est<-predict(m,df$x)
  plot(x,y)
  lines(x,y_est)
  summary(m)
```

ملحق (B)

برنامج لتدريب الشبكة العصبية والتنبؤ بمعدلات المواليد باستخدام حزمMATLAB.

```
%برنامج لتدريب الشبكة العصبية والتنبؤ بمعدلات المواليد
باستخدام حزمMATLAB.

 x= [1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33
34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49
50 51 52 53 54 55 56 57 58 59 60 ];

y = [43.1 44.1 44.5 43 42.3 41.7 41.2 39.2 38.2
37 35.1 35.1 34.4 35.7 35.7 36 36.4 37.3 38.8
40.8 37.5 37 36.2 36.8 36.6 39.8 38.7 38.8 37.8
33.2 31.7 30 26.9 28.1 27.7 27.9 28.3 27.5 27.5
27 27.4 26.7 26.5 26.1 25.7 25.5 25.7 26.5 27.3
28.8 28.7 30.3 31.9 31 31.3 30.2 28,6 26.8 24,5
23.4];
 n=size(y'); h= 3 ;
                        % انشاء وتدريب الشبكة.
                                  ..............
          net = fitnet(2,'trainlm');
           % net.trainParam.showWindow =0 ;

  [net,tr]= train(net, x,y);
                % التنبؤ بقيم معدلات المواليد.



            y0=sim(net ,x);
```