طريقة مقترحة لتحليل الانحدار غير الخطى المتعدد باستخدام الذكاء الاصطناعى

ا.م.د/ عبد الوهاب السيد حجاج
رئيس قسم الإحصاء – وكيل كلية التجارة – جامعة الازهر
MAIL:Wahabsat1@yahoo.com
Mobile: 0020-01011604005

د/ محمد عبدالسلام العجمى
مدرس بقسم الإحصاء – كلية التجارة – جامعة الازهر
MAIL:dr.muhammad_alagamy@azhar.edu.eg
Mobile: 0020-01024255565

ملخص البحث

تنشآ مشكلة هذا البحث من انه قد تواجه الطرق التقليدية صعوبات عديدة في صياغة النموذج الملائم لبيانات معينة ، بعض هذه الصعوبات صعوبات رياضية ، وبعضها الحاجة إلى خبرة كبيرة لاختيار النموذج الملائم للبيانات , أنه في حالة انحدار المشاهدات الحالية Y_t على الأخطاء السابقة يكون لدالة الإمكان شكلا معقدا بسبب أن الأخطاء لها علاقة غير خطية في المعلمات مما يصعب معه التعبير عن دالة التوزيع الاحتمالي لبعض أو كل معالم النموذج في شكل قياسي أو باستخدام أحد التوزيعات المعروفة.

في هذا البحث تم تقديم طريقة مقترحة لتحليل الانحدار غير الخطى المتعدد باستخدام أحد طرق الذكاء الاصطناعي وهي طريقة الشبكات العصبية الاصطناعية. قام الباحث باستخدام بيانات دخل ١٠٠٠ اسرة والتي تم الحصول عليها من خلال بحث الدخل والانفاق والاستهلاك لعام ٢٠١٧ / ٢٠١٨والذى يجريه الجهاز المركزى للتعبئة العامة والاحصاء , ولقد قام الباحث بتقدير معالم نموذج الانحدار غير الخطى المتعدد فكان افضل نموذج هو النموذج اللوغاريتمى

قام الباحث باستخدام برنامج MATLAB فى تدرب الشبكة العصبية الاصطناعية وذلك من خلال انشاء شبكة أمامية بتغذية مرتدة وتم تقدير المعالم والتحقق من صلاحية النموذج باستخدام أسلوب التغذية الخلفية Back propagation algorithm

وتم حساب القيم التنبؤية لإجمالي الدخل للأسر بجانب العشرة مشاهدات المستقبلية عن طريق دوال MATLAB للشبكات العصبية وكذلك برنامج SPSS لأسلوب الانحدار غير الخطى المتعدد ,كما تم حساب دقة هذه التنبؤات بواسطة ثلاثة مقاييس إحصائية وهى (MSE, MAD, MAPE) وكانت نتيجة المقارنة هى افضلية اسلوب الانحدار غير الخطى المتعدد عن اسلوب الشبكات العصبية الاصطناعية للتنبؤ بقيمة الدخل للأسرة.

كلمات مفتاحية:

الذكاء الاصطناعي Artificial Inelegant– الشبكات العصبية الاصطناعية – الانحدار الخطى – الانحدار غير الخطى المتعدد – دخل الاسرة – مستوى المعيشة .

(١) مقدمة

يعد تحليل الانحدار أحد أهم الأساليب التي تستخدم على نطاق واسع في جميع مجالات المعرفة والعلم حيث يمكن باستخدامه التنبؤ بدرجة أو بقيمة المتغير التابع من خلال معرفة قيم المتغيرات المستقلة. بمعنى آخر معرفة أكثر المتغيرات المستقلة تأثيراً في المتغير التابع.

وهو امتداد للنموذج الانحدار الخطى البسيط ففي حين أن نموذج الانحدار الخطى البسيط كان يعتمد على متغيرين متغير تابع والآخر متغير مستقل، فإن النموذج العام قد يتضمن عدد من المتغيرات قد يكون من بينها متغير تابع واحد والعديد من المتغيرات المستقلة. وتعد مشكلة تحديد النموذج المستخدم في عملية التنبؤ أو التقدير باستخدام نماذج الانحدار من أهم المشكلات التي تواجه الباحثين

حيث تتداخل العديد من العوامل والمتغيرات في تحديد اختيار النموذج الملائم للتحليل وبالتالي فيجب تصنيف نماذج الانحدار تبعا لهذه العوامل والمتغيرات التي تؤثر فيها، ومنها: (عدد المتغيرات، شكل العلاقة بين المتغيرات، مستوى القياس للمتغيرات، طريقة جمع البيانات).

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 اصبح 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 الاصطناعيArtificial 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. ومن لغاته هي :لغة Lisp: وهى لغة معالجة اللوائح (List Processing Language), ولغة Prolog :وهى لغة البرمجة بالمنطق (Programming in Logic) , وهذه اللغات تتيح للمبرمج إمكانيات كبيرة في كتابة البرامج . ويشمل الذكاء الاصطناعي فروعا عديدة من أهمها : النظم الخبيرة Expert Systems, والتنقيب في البيانات Data Mining ,

والخوارزميات الجينية Genetic algorithms, والشبكات العصبية Neural Networks.

(١-١) مشكلة البحث

تكمن مشكلة البحث أنه في حالة تحليل الانحدار في وجود علاقة غير خطية يكون لدالة الإمكان شكلا معقدا بسبب أن الأخطاء لها علاقة غير خطية في المعلمات مما يصعب معه تحديد شكل التوزيع الاحتمالي لبعض أو كل معالم النموذج في شكل قياسي أو باستخدام أحد التوزيعات المعروفة وأيضا عند استخدام أسلوب بييز حيث يصعب إيجاد " مصفوفة التباين والتغاير " ومحدد مصفوفة التباين والتغاير, واختيار دالة احتمال قبلية مناسبة لها, خاصة مع زيادة حجم العينة, ولتجنب مثل هذه المشاكل نحتاج إلي طريقة جديدة للتنبؤ بخلاف الطرق الإحصائية التقليدية (الطرق والاساليب المعلمية), وهذه الطريقة هي الشبكات العصبية الاصطناعية, حيث تمتاز هذه الطريقة بأنها تعمل بصورة آلية (الطرق والاساليب اللامعلمية), لاختيار النموذج المناسب والتنبؤ بقيم الظاهرة فى المستقبل .

(١-٢) هدف البحث

في هذا البحث نقدم طريقة مقترحة لتحليل الانحدار غير الخطى المتعدد باستخدام أحد طرق الذكاء الاصطناعي وهي طريقة الشبكات العصبية الاصطناعية . وتُعد الشبكات العصبية الاصطناعية احد الطرق الحديثة نسبيا في التنبؤ بالانحدار المتعدد, بالإضافة إلى قدرتها على معالجة البيانات دون شروط مسبقة

(١-٣) أهمية البحث

تأتى أهمية البحث في استخدامه لأحد طرق الذكاء الاصطناعي وهي طريقة الشبكات العصبية في تحليل الانحدار المتعدد، حيث أنها تمتاز عن غيرها من الأساليب الإحصائية التقليدية بمزايا منها: أنها تفيد في الحصول على تنبؤات جيدة , حيث أنها تعطى مجموع مربعات أخطاء غير خطية أقل , كما أن طريقة الشبكات العصبية الاصطناعية (ANN) تحتاج عينات أصغر نسبيا في اختبار مصداقية

النماذج , ومن مزاياها أيضا أنه يتم التنبؤ فيها بصورة آلية وأنها تصلح في حالة النماذج الخطية وغير الخطية, كما يمكن استخدامها لتقييم قدرة الأساليب الإحصائية الأخرى في تقدير النماذج واستخدامها في التنبؤ . ويمكن تلخيص أهمية البحث فى :

١- **أهمية تطبيقية** : و تتركز الأهمية التطبيقية لهذه الدراسة في أنها تساعد المؤسسات الكبيرة والدول في وضع خططها المستقبلية بناء على قراءة لتطور أحوال دخول الاسر وانفاقهم المعيشى .

٢- **أهمية إحصائية** : تكمن الأهمية الإحصائية لهذه الدراسة في استخدام الطرق والاساليب اللامعلمية, مثل الشبكات العصبية الاصطناعية كأحد طرق الذكاء الاصطناعى للتنبؤ بحجم الدخل للأسر على المدى البعيد باستخدام بيانات لمدة زمنية طويلة .

(١-٤) حدود الدراسة :

اولا الحدود الزمنية : لقد قام الباحث باستخدام بيانات تم الحصول عليها من خلال بحث الدخل والانفاق والاستهلاك لعام 2017/2018 والذى يجريه الجهاز المركزى للتعبئة العامة والاحصاء المصرى لعينة من ١٠٠٠ اسرة.

ثانيا الحدود المكانية :

لقد تم التنبؤ بقيم دخول الاسر داخل جمهورية مصر العربية باستخدام الطريقة التقليدية متمثلة فى نموذج الانحدار غير الخطى المتعدد وكذلك طريقة الشبكات العصبية الاصطناعية كأسلوب من اساليب الذكاء الاصطناعى كطرقة حديث فى التنبؤ .

(١-٥) دراسات سابقة في مجال الذكاء الاصطناعي والانحدار المتعدد :

تعرض بعض الباحثين والإحصائيين لتحليل الانحدار المتعدد باستخدام الشبكات العصبية :

فقد جاءت دراسة ندوى خزعل رشاد ، عزة حازم زكى – مجلة تنمية الرافدين العدد ٩٩ مجلد ٣٢ لسنة ٢٠١٠ ، حيث تعرضا إلى تحليل الانحدار المتعدد باستخدام الشبكات العصبية ومقارنته مع أسلوب تحليل الانحدار التدريجي Stepwise Regression وتم التوصل إلى أن الشبكات العصبية ذات دقة أكثر من طريقة الانحدار الخطي المتعدد في كشف المتغيرات ذات الفروقات غير المعنوية .

وفي دراسة زكريا يحيي الجمال ، عمر صابر – المجلة العراقية للعلوم الإحصائية العدد ٢١ لسنة ٢٠١٢ ، حيث استخدما أسلوب مهجن بين الشبكات العصبية الاصطناعية ونموذج الانحدار الخطي إذ تم توضيح آلية عمل هذا الأسلوب المهجن ومقارنته مع نماذج الشبكات العصبية ونموذج الانحدار الخطي لمعرفة مدى كفاءة هذا الأسلوب وتم التوصل إلى كفاءة هذا الأسلوب المهجن .

وفي دراسة عائدة يونس محمد المراد – المجلة العراقية للعلوم الإحصائية العدد ٢١ لسنة ٢٠١٢ ، حيث قاما بتحليل الانحدار باستخدام الطريقتين الكلاسيكية والتقنيات الذكائية وإجراء المقارنة بينهما وتم التوصل إلى أن الشبكات العصبية أفضل في نتائجها .

وفي دراسة دربال ٢٠١٤ تم مقارنة نماذج التنبؤ الخطية وغير الخطية للتنبؤ بمؤشر سوق دبي المالي وذلك بالاعتماد على قاعدة بيانات يومية للفترة ٢٠٠٦/٢/٢٢ إلى٢٠١٤/١/٣٠ وتوصلت الدراسة إلى أن نموذج الشبكات

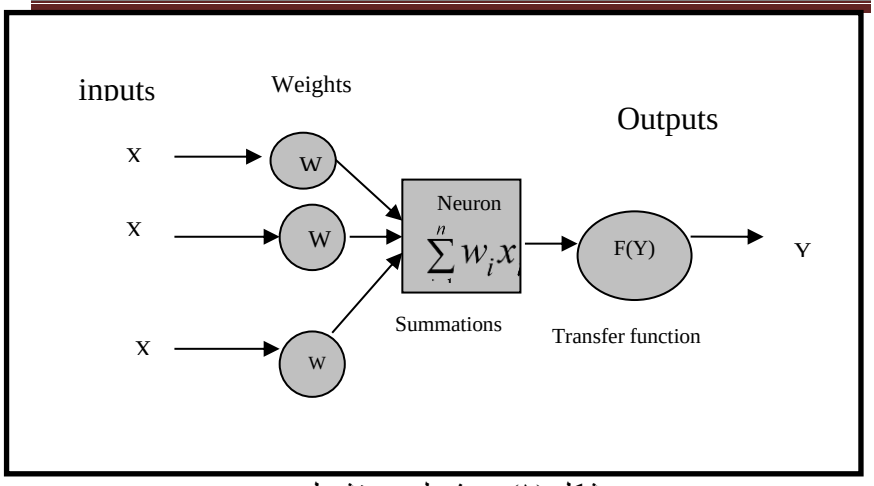
العصبية الاصطناعية ANN لديه قدرة أكبر على التنبؤ مقارنة بنموذج الانحدار الذاتي المشروط بعدم تجانس الأخطاء المعمم GARCH.

وقد هدفت دراسة مستور , عبد الرحيم (2017) إلى معالجة مشكلة التداخل الخطي المتعدد بين المتغيرات المستقلة عن طريق استخدام أسلوب المكونات الرئيسية

وتهدف هذه الدراسة الحالية إلى تقديم طريقة جديدة ومؤثرة للتنبؤ باستخدام تقنية الشبكات العصبية الاصطناعية (ANN) Artificial Neural Network , والمقارنة بين الطريقة المقترحة والطريقة التقليدية للانحدار المتعدد غير الخطى.

(١-٦) الشبكات العصبية الاصطناعية

تعد الشبكات العصبية الاصطناعية (ANN) Neural Networks 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 Output Signals(Y) .

وتتشابه الشبكات العصبية الاصطناعية ANN مع العقل البشري في أنها تكتسب المعرفة بالتدريب وتخزن هذه المعرفة باستخدام قوى داخل العقد تسمى

الأوزان , وهناك تشابه عصبي حيوي مما يعطي الفرصة لعلماء البيولوجيا في الاعتماد على ANN لفهم تطور الظواهر الحيوية .

وتستخدم الشبكات العصبية في الكثير من المجالات و التطبيقات منها : معالجة الإشارة Signal Processing ، السيطرة Control ، تمييز الأنماط Pattern Recognition ، الطب Medicine ، إنتاج الكلام Speech Production ، تمييز الكلام Speech Recognition [علام زكي ،٢٠٠٠ ، ص١٥-١٩]. وخلال هذا البحث سنبين كيف يمكن استخدام أحد فروع الذكاء الاصطناعي وهو الشبكات العصبية فى تحليل الانحدار المتعدد.

(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 (1992) Zurada).

و يتم 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 Single-Layer Feed forward Network , (٢) شبكة متعددة الطبقات ذات التغذية الأمامية Multi-Layer Feed forward Network , (٣) شبكة متعددة الطبقات ذات التغذية المرتدة Multi-Layer Recurrent Network .(انظر: Demuth & Beale 1998) و سنوضح كل نوع فيما يلي :

١ـ 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 :

الخطـوة الأولـى: التحديد العشوائي للقيم الابتدائية لـلأوزان $w_1,w_2,\ldots\ldots,w_n$, وللقيمة الحدية Threshold Value (θ) وذلك في المدى $[-0.5,0.5]$.

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:

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

X : صافى المدخلات الموزونة للخلية العصبية

Wi : الأوزان النسبية لعقد الاتصال ما بين الطبقات

xi : قيمة المدخلات i

n : عدد مدخلات الخلية العصبية

يتم تنشيط المدرك Perceptron عن طريق تطبيق المدخلات X1, X2,... Xn , والمخرجات المرغوب بها T(k) , وحساب المخرجات الفعلية عند التكرار k كما يلي :

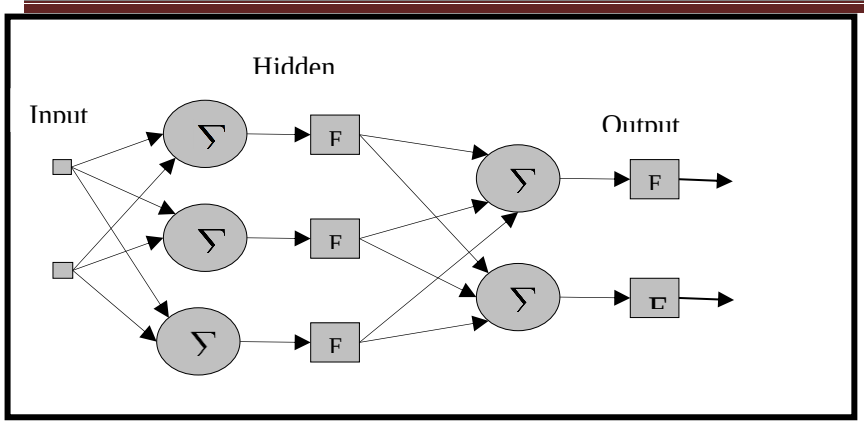
$$y(k) = step[\sum_{i=1}^{n} X_i(k)W_i(k) - \theta] \qquad (2)$$

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) .

١- متعددة الطبقات ذات التغذية الأمامية

تتكون الشبكة متعددة الطبقات من واحد أو أكثر من الطبقات (أو المستويات) من عقد الاتصال Nodes تسمى 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 :(2007) Hijazi



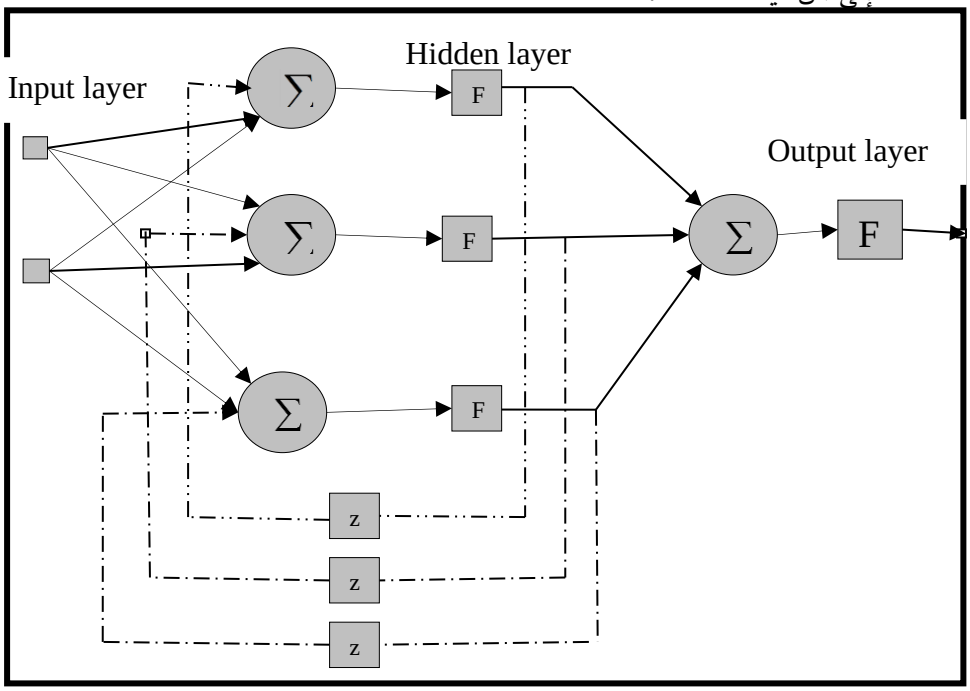
شكل(٢) الشبكة متعددة الطبقات ذات تغذية أمامية

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 (1989) Cybenko) .

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 , وهذا 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 الخطأ 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 غير الخطي المتعدد Non-Linear Multiple Regression

في الظواهر الاقتصادية غالباً ما تكون المتغيرات محل الدراسة – وفقا لمعادلة المنحنى على شكل الانتشار أو نتيجة لخبرة سابقة – تعبر عن علاقات غير خطية كعلاقة معدل التكلفة بكمية الإنتاج:

$$Y = a + b X + c X^2 \qquad (3)$$

وهي معادلة من الدرجة الثانية حيث a, b , c قيم ثابتة، وتعرف بعلاقة الانحدار التربيعي.

والهدف الحصول على أفضل خط انحدار سواء كان خطي أو غير خطي.

والأفضلية بدءاً للخط المستقيم وإلا نقوم بإضافة تربيع أو تكعيب للمتغير المستقل X. الا إذا كنا نعلم مسبقاً بأن المتغيرات لدينا ذات علاقة تربيعية فنبدأ باحتسابها.

ونبدأ هنا بالانحدار غير الخطي البسيط وصولا للمتعدد.

الانحدار غير الخطي البسيط Simple Non-Linear Regression

من المعلوم ان الانحدار الخطي البسيط يعبر عنه بالمعادلة:

$$Y = a + b X \qquad (4)$$

نحسب a, b من:

$$b = \frac{n\sum XY - \sum X \sum Y}{n\sum (X)^2 - (\sum X)^2}$$

$$a = \frac{\sum Y - b\sum X}{n}$$

ويختلف غير الخطي البسيط عن الخطي البسيط فيما يلي:

المعامل a الثابت ليس حداً (كما في: $Y = a + b X$) ولكنه قد يظهر بالصورة:

$$Y = aX^b \qquad (5)$$

ويمكن أيضا تحويلها لخطية (لوغاريتم):

$$Y = a + \ln X \qquad (6)$$

كما يمكن تحويلها لخطية بتغيير تعريف المتغيرات، فمثلاً :

$$Y = aXb \qquad (7)$$

بأخذ لوغاريتم الطريفين نحصل على:

$$\ln Y = \text{Lina} + b \ln X \qquad (8)$$

وذلك حسب قوانين الأسس وبوضع:

$$\ln Y = Y \text{ , } \ln a = a \text{ , } \ln X = X$$

نحصل على المعادلة المرادفة:

$$Y = a + bX \qquad (9)$$

وبالتالي نحسب قيم a , b .

ويمكن تقدير المتغير بإعادة تعريف المتغيرات كما في المعادلة:

$$Y = a + b/X$$

$$Y = a + b(1/x) \qquad \text{اى :}$$

وهنا نستخدم نفس القوانين السابقة بوضع $1/X$ بدلا من X .

ويمكن تقدير المتغير بإعادة تعريف المتغيرات كالمعادلة:

$$Y = a + b \ln X \qquad (10)$$

وتعرف بالمعادلة نصف اللوغاريتمية باستبدال lnX بدل من $1/X$ في الصيغ السابقة أي:

$$b = \frac{n\sum \text{YLinX} - \sum \text{LinX} \sum Y}{n(\sum \text{LinX}^2) - (\sum \text{LinX})^2}$$

$$a = \frac{\sum Y - b\sum \text{LinX}}{n}$$

الانحدار غير الخطي المتعدد Non-Linear Multiple Regression

سبق أن ذكرنا في الانحدار غير الخطي البسيط بوجود أنواع من الانحدار غير الخطي والتي تضم العنصر X^2 أو X^3 أي إضافة أحد العنصرين X^2 أو X^3

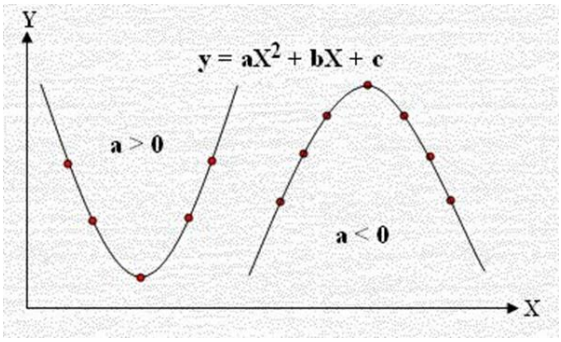
لمعادلة الانحدار غير الخطي البسيط ونحصل على معادلة الانحدار التربيعي ، ومعادلة الانحدار التكعيبي.

اولا: معادلة الانحدار التربيعي Quadratic Regression Equations

هذه المعادلة تعتبر من أبسط معادلات الانحدار غير الخطي المتعدد وهي معادلة من الدرجة الثانية ونشأت من إضافة العنصر X^2 لمعادلة الانحدار الخطي البسيط وتأخذ الصورة الآتية:

$$y = a + b X + c X^2 \qquad (11)$$

والمنحنى قطع مكافئ مفتوح من أعلى (معامل X2 موجب $a > 0$) ورأسه نقطة صغرى أو مفتوح من أسفل، (معامل X2 سالب $a < 0$) ورأسه نقطة عظمى والرأس في الحالتين له إحداثي سيني = $(b \div 2a)$ – كما بالشكل:



ولتحويل معادلة الانحدار التربيعي لمعادلة خطية نفرض الآتي:

$$X_1 = X \quad , \quad X_2 = X^2$$

فتكون لدينا الصيغة:

$$y = a + b X_1 + c X_2 \qquad (12)$$

وبالتالي يمكن إيجاد قيم الثوابت باستخدام طريقة المربعات الصغرى وبنفس الطريقة التي استخدمت في معادلة الانحدار الخطي المتعدد لحل المعادلات الثلاث التالية:

$$\sum Y = na + b\sum X_1 + c \sum X_2$$

$$\sum X_1 Y = a\sum X_1 + b\sum X_1^2 + c \sum X_1 X_2$$

$$\sum X_2 Y = a\sum X_2 + b\sum X_1 X_2 + c\sum X_2^2$$

معادلة الانحدار التكعيبي Cubic Regression Equation

هذه المعادلة من الدرجة الثالثة وهي امتداد للمعادلة التربيعية والحاجة للقوة الثالثة أي الدرجة الثالثة أي إضافة X3 للمعادلة التربيعية لضبط البيانات الموجودة والصورة العامة لمعادلة الدرجة الثالثة (التكعيبية) هي:

$$Y = a + bX + cX^2 + dX^3 \qquad (13)$$

نحول هذه المعادلة لصورة خطية بفرض:

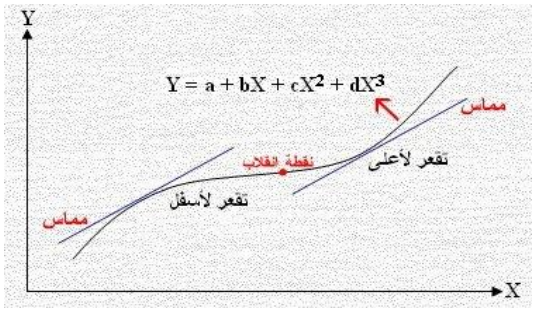
$$X = X_1 \ , \ X^2 = X_2 \ , \ X^3 = X_3$$

فتكون المعادلة الخطية بعد الفرضية هذه هي:

$$Y = a + bX_1 + cX_2 + dX \qquad (14)$$

ويمكن حساب قيم الثوابت هنا a , b , c , d جبريا وهو أمر ليس بالسهل ولذا نستخدم الحاسب الآلي من خلال MINITAB , SPSS , ...

والتمثيل البياني للمعادلة التكعيبية يبينها الشكل التالي



حيث أن:

التقعر لأعلى: المنحنى يقع فوق مماساته.

التقعر لأسفل: المنحنى يقع أسفل مماساته.

نقطة الانقلاب: النقطة التي يتحول عندها المنحنى من تقعره لأسفل إلى تقعره لأعلى أو هي النقطة التي تكون عندها المشتقة الثانية تساوي صفر.

والشكل يبين شكل المعادلة:

$$Y = a + bX + cX^2 + dX^3 \qquad (15)$$

وهي الصورة العامة لمعادلة الدرجة الثالثة.

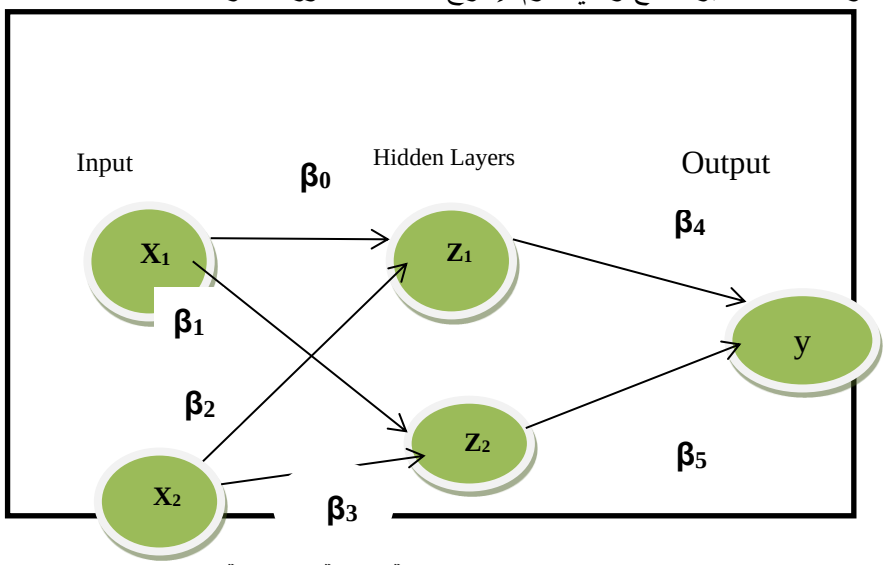
(٨-١) الشبكات العصبية كنموذج انحدار غير خطى:

من الممكن اعتبار الشبكات العصبية الاصطناعية كنموذج انحدار غير خطى يتكون من المتغيرات المدخلة وتتمثل في القيم الماضية $y_{t+1},y_{t+2},\ldots,y_{t+p}$ أو تحويلاتها و قيم متغير خارجي $x_{t+1},x_{t+2},\ldots,x_{t+h}$ والأخطاء :

وكما فى الشكل (٤) تمثل X متغيرات مستقلة داخلة في البرنامج " Input " ,

وتمثل Z 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 :

$$Y = X\beta + \varepsilon \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (16)$$

حيث Y: متجه من الدرجة $(n\times1)$ تمثل مشاهدات عن المتغير التابع

X: مصفوفة من الدرجة (n×k) تمثل مشاهدات المتغيرات المفسرة وعددها (k − 1)

β : متجه عمود يمثل معالم المجتمع وعددها k .

ε : متجه عمود من الدرجة (n × 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 (العينة sample) , 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 :

$$Z = g(X,W) + \varepsilon \qquad (17)$$

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 :

$$Y = Z \cdot W_2 + \varepsilon_2 \qquad (18)$$

حيث :

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

$$Z = X.W1 + \varepsilon_1 \qquad (19)$$

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

حيث :

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

وبالتعويض عن Z في الدالة Y نحصل على :

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

$$Y = X W_2 W_1 + \varepsilon_1 W_2 + \varepsilon_2 \qquad (20)$$

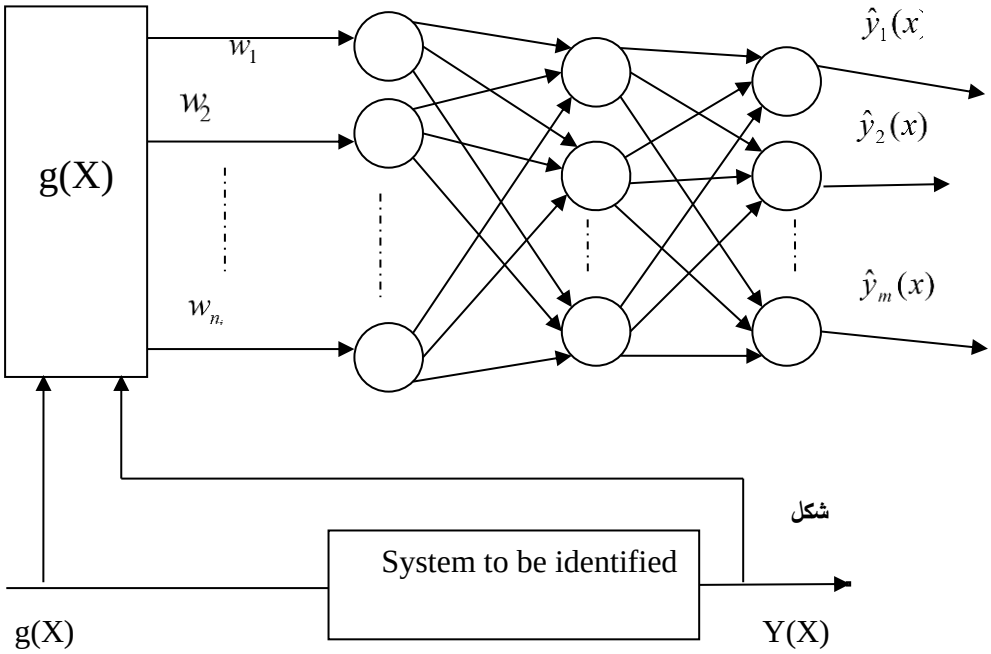
وبوضع : $\beta = W_2 W_1 \quad , \quad \varepsilon = \varepsilon_1 W_2 + \varepsilon_2$

نستنتج أن :

$$Y = X \beta + \varepsilon \qquad \ldots\ldots\ldots \qquad (21)$$

وبالتالي نكون قد حصلنا على النموذج العام للانحدار المتعدد .ومن الممكن إضافة أكثر من طبقة مخفية وسنحصل على نموذج مشابه للنموذج (1-4) 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)



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

(١-٩) الدراسة تطبيقية

مقدمة :

تهتم الكثير من البلدان بتقليص التباين فى توزيع الدخول بين مجتمعاتها وتحقيق التوازن فى توزيع الثروة والقضاء على الفقر لهذا اهتمت غالبية الدول بين فترة واخرى بإجراء مسوح لميزانية الاسرة (الدخل والانفاق) والسعي لعمل دراسات وابحاث تتناول هذا الموضوع .

ويعد بحث الدخل والانفاق والاستهلاك من اهم الابحاث التي يجريها الجهاز المركزي للتعبئة العامة والإحصاء، حيث يوفر كم هايل من البيانان والمعلومات التي يتم

استخدامها في قياس مستوى معيشة الاسرة والافراد، وكذا توفير قواعد لبيانات نسب الفقر وتحديد الفئات المستهدفة لبرامج الحماية المجتمعية المختلفة.

وقد ادركت مصر اهمية هذا البحث منذ فترة طويله ، حيث قامت بعمل اول بحث عام ١٩٥٨\ ١٩٥٩تلاه العديد من الأبحاث كان حوالي ١٤ عشر بحثا بعدها في سلسلة الأبحاث التي قام الجهاز بتنفيذها.

وتتألف عضوية **الأسرة المعيشية** من جميع الأشخاص المقيمين عادة في الأسرة المعيشية، حيث يتم تعريف الإقامة المعتادة بأسلوب يتمشى مع الأحكام الواردة في التنقيح الأخير لمبادئ وتوصيات تعدادات السكان والمساكن للأمم المتحدة. ويمكن استخدام فترة يبلغ حدها الأدنى ستة أشهر كمعيار من معايير تحديد الإقامة المعتادة.

رب الاسرة: معيار اختيار شخص وحيد المعترف به كرب الأسرة هو الشخص الذي يتخذ القرارات الهامة والشخص الأكبر سناً بين الموجودين وما إلى ذلك.

إنفاق الأسرة المعيشية: يعّرف على أنه مجموع الإنفاق الاستهلاكي للأسرة المعيشية والنفقات غير الاستهلاكية للأسرة المعيشية. والنفقات غير الاستهلاكية هي النفقات التي تتحملها أسرة معيشية فيما يتصل بالتحويلات الإلزامية وشبه الإلزامية إلى الحكومة والمؤسسات غير الهادفة للربح والأسر المعيشية الأخرى دون حيازة أي سلع أو خدمات مقابل إشباع احتياجات أفرادها .ويمثل إنفاق الأسرة المعيشية مجموع النفقات التي يتعين على الأسرة المعيشية أن تتحملها لإشباع احتياجاتها واستيفاء التزاماتها "القانونية."

لتقدير الإنفاق الاستهلاكي للمساكن التي يشغلها مالكها يتم جمع البيانات بشأن :

١- سمات المسكن (سنة وحجم ونوع البناء والمرافق وتكاليف الصيانة والإصلاح ووضع الجوار

٢- إيجارات المساكن المؤجرة (من استقصاء أو من مصادر أخرى)

٣- والقيمة السوقية للمساكن

٤- تكاليف السكن التي يدفعها الملاك في العادة عن جميع المساكن

٥- مدة الاستخدام بالنسبة للمنازل المخصصة للإجازات أو لعطلة نهاية الأسبوع

دخل الأسر المعيشية يتكون من جميع الإيرادات التي تتلقاها الأسرة المعيشية أو أفراد الأسرة المعيشية، نقدا أو عينا أو خدمات، سنويا أو على فترات أقرب، ولكنه يستبعد الأرباح الطارئة وغيرها من مثل هذه الإيرادات غير المنتظمة والعارضة في العادة.

وبالنسبة لهذا البحث فلقد قام الباحث باستخدام بيانات تم الحصول عليها من خلال بحث الدخل والانفاق والاستهلاك لعام2017/2018 والذى يجريه الجهاز المركزى للتعبئة العامة والاحصاء المصرى لعينة من ١٠٠٠ اسرة وتم التنبؤ بقيم دخول الاسر باستخدام الطريقة التقليدية متمثلة فى نموذج الانحدار غير الخطى المتعدد وكذلك طريقة الشبكات العصبية الاصطناعية كأسلوب من اساليب الذكاء الاصطناعى كطرقة حديث فى التنبؤ وفيما يلى شرح لتطبيق كلى الطريقتين :

(١-٩-١) النماذج التقليدية:

قام الباحث باستخدام برنامج SPSS بتقدير معالم نموذج الانحدار غير الخطى المتعدد لبيانات دخل الاسرة فكان افضل نموذج كما يلى :

$$\text{Log } \mu_y = 1.13 + (-0.025)\, x_1 + (0.003)\, x_2 + (-0.007)\, x_3 + (0.006)\, x_4 \qquad (22)$$

حيث ان :

μ_{y1} : القيمة المتوقعة للمتغير التابع (اجمالي الدخل)

x_1 :المتغير المستقل (حجم الاسرة)

x_2 : المتغير المستقل (السن بالسنوات)

x_3 : المتغير المستقل (الانفاق على السكن)

x_4 : المتغير المستقل(الانفاق على الطعام والشراب)

(١-٩-٢) طريقة الشبكات العصبية الاصطناعية :

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.

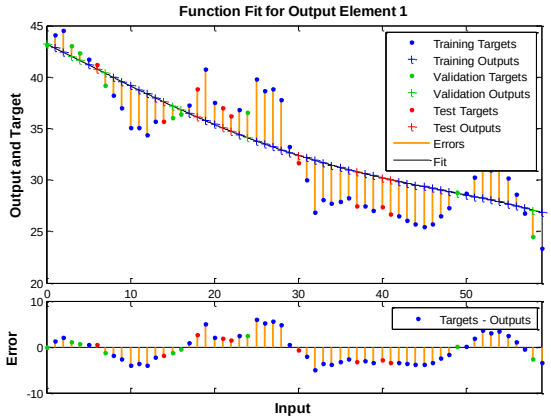
إنشاء الشبكة network creation :

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

```
net = fitnet(5,'trainlm')                    (23)
```

(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).



شكل (٦) يوضح الاداء لشبكة التغذية الأمامية .

تقدير معالم النموذج Estimation of Model parameters

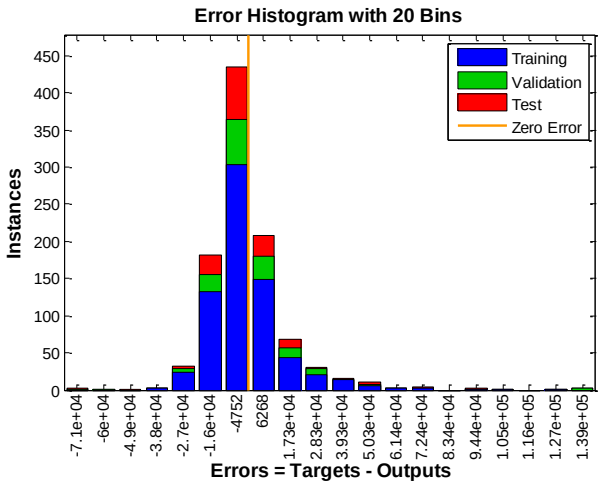
يقصد به تقدير أوزان الشبكة (معالم النموذج) باستخدام أسلوب الشبكات العصبية الاصطناعية , و لكي يتم ذلك يلزم جعل متوسط مربعات الخطأ (MSE) اقل ما يمكن . و من الممكن الحصول على تقدير أوزان الشبكة (معالم النموذج) عن طريق تدريب الشبكة باستخدام مجموعة بيانات التدريب , حيث تمثل هذه الأوزان المعلومات الأولية التي 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 Algorithm Back Propagation التي تستخدم في تدريب الشبكات العصبية متعددة الطبقات ذات التغذية الأمامية وغير الخطية ، وتتم عملية تدريب الشبكة داخل برنامج MATLAB بالتسلسل التالي :

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

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) .

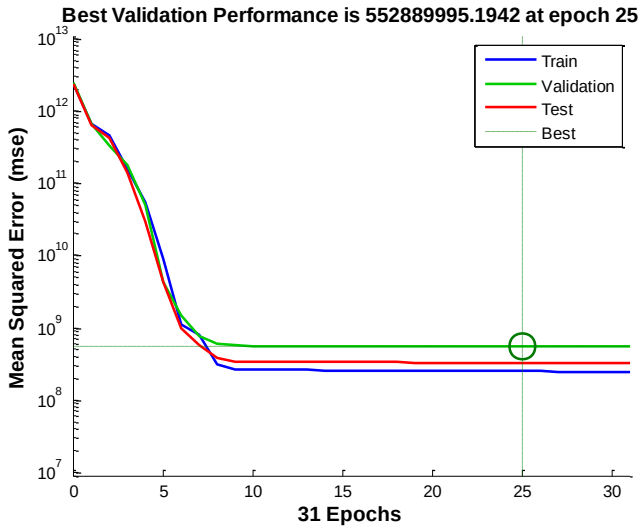
شكل (٧) يوضح المدرج التكرارى للأخطاء .



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

بالإضافة إلى معايير تشمل عدد مرات إجراء تعديل الأوزان (معالم النموذج) وتسمى Epoch والتي يمكن التحكم بها في النموذج (وكما فى الشكل التالى تم اجراء 31 Epoch للوصول الى افضل اداء للشبكة لتحقيق اقل اخطاء ممكنة والذى تحقق عند25 Epoch) .



شكل (٨) يوضح افضل اداء للشبكة .

استخدام النموذج فى 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.

ويتم دراسة هذه المرحلة كما يلي : عند التنبؤ بقيم الانحدار المتعدد باستخدام الشبكات العصبية الاصطناعية بالاستعانة ببرنامج MATLAB تم استخدام الدالة التالية :

$$Y = sim(net, x) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (23)$$

حيث : Y : تمثل القيم التنبؤية باستخدام أسلوب الشبكات العصبية الاصطناعية .

net : تمثل الشبكة المدربة .

X : تمثل البيانات التي تم استخدامها لتدريب الشبكة

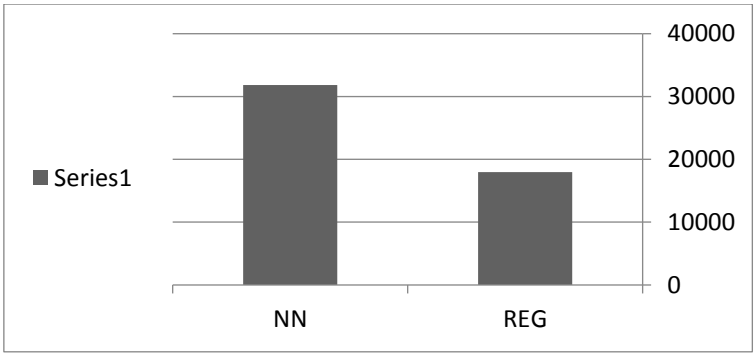
(١-٩-٣) المقارنة بين نتائج الانحدار المتعدد غير الخطى و الشبكات العصبية الاصطناعية:

بعد حساب القيم التنبؤية لإجمالي الدخل للأسر بجانب العشرة مشاهدات المستقبلية عن طريق دوال MATLAB المشار إليها سابقا للشبكات العصبية وكذلك برنامج SPSS لأسلوب الانحدار غير الخطى المتعدد , ولقد تم حساب دقة هذه التنبؤات بواسطة ثلاثة مقاييس إحصائية وهى (MSE, MAD, MAPE) والنتائج التى تم الحصول عليها تم وضعها في الجدول التالى :

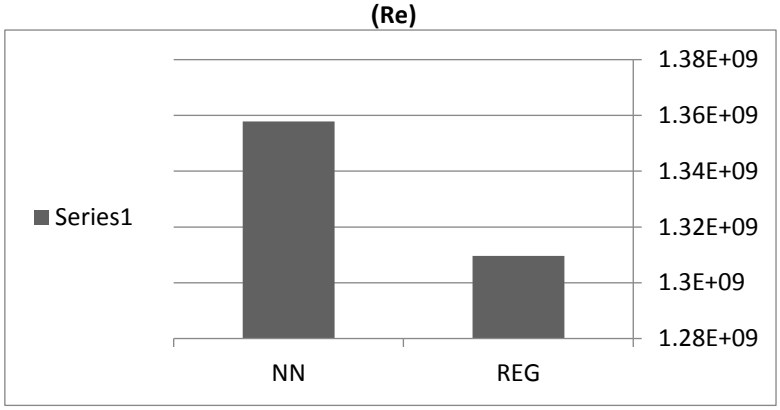
Year	Y	Y(predict)		MAD		MSE		MAEP	
		NN	REG	NN	REG	NN	REG	NN	REG
1	172512.5	87618.5	61628.7	80767	110881.3	6523308289	12294658254	0.461	0.647
2	47699	65827.7	47263.9	16951	435.09	287336401	189303.3081	0.353	0.001
3	51719	78471	50860	22143	858.94	490312449	737777.9236	0.428	0.0166
4	54306	76987	42344.7	26842	11961.23	720492964	143071023.1	0.4942	0.2202
5	71880	6838.9	70118.6	18511	1761.38	342657121	3102459.504	0.2575	0.0245
6	43737	7032.5	55961.4	33332	12224.43	1111022224	149436688.8	0.7621	0.2794
7	62543	6981.1	67260.2	28542	4717.2	814645764	22251975.84	0.4563	0.0754
8	58527	104713.2	75139.2	47743	16612.2	2279394049	275965188.8	0.8157	0.2838
9	57453.5	87363.6	48776.5	27535	8677.42	758176225	75297617.86	0.4792	0.1510
10	100473.5	111023.3	88996.1	15830	11473.89	250588900	131650151.7	0.1575	0.1142
Average				31819.6	17960.31	1357793439	1309636044	46.7	18.138

جدول (١) نتائج (MSE,MAD, MAPE) لتنبؤات الشبكات العصبية (NN) وتنبؤات الانحدار غير الخطى المتعدد (Re)

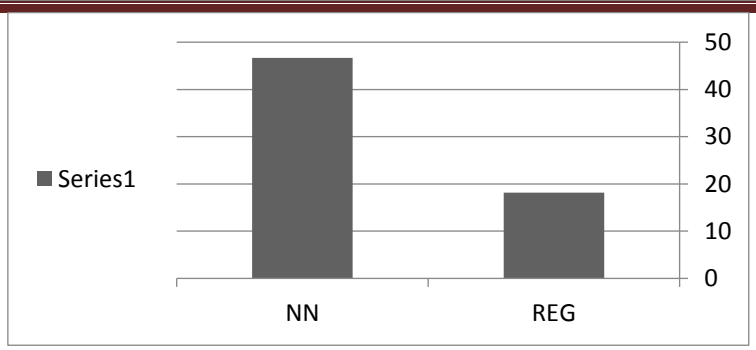
من خلال الجدول السابق يتضح لنا ان اغلب قيم MAD ,MSE, MAPE لتنبؤات اسلوب الانحدار المتعدد غير الخطى كانت اقل من الشبكات العصبية الاصطناعية, كذلك كان المتوسط العام لهذه المقاييس فى حالة كل المشاهدات (الف اسرة) لتنبؤات اسلوب الانحدار المتعدد غير الخطى اقل من الشبكات العصبية الاصطناعية , مما يعنى افضلية اسلوب الانحدار غير الخطى المتعدد عن اسلوب الشبكات العصبية الاصطناعية للتنبؤ بقيمة الدخل للأسرة.



شكل (٩)
يوضح متوسط قيم MAD لتنبؤات الشبكات العصبية (NN) وتنبؤات الانحدارغير الخطى المتعدد (Re)



شكل (١٠)
يوضح متوسط قيم MSE لتنبؤات الشبكات العصبية (NN) وتنبؤات الانحدار غير الخطى المتعدد (Re)



شكل (١١)
يوضح متوسط قيم MAPE لتنبؤات الشبكات العصبية (NN) وتنبؤات الانحدار غير الخطى المتعدد (Re)

(١-١٠) النتائج والتوصيات

أولاً: النتائج

١- في هذا البحث تم تقديم طريقة مقترحة لتحليل الانحدار غير الخطى المتعدد باستخدام أحد طرق الذكاء الاصطناعي وهي طريقة الشبكات العصبية الاصطناعية .

٢- قام الباحث باستخدام بيانات دخل ١٠٠٠ اسرة والتى تم الحصول عليها من خلال بحث الدخل والانفاق والاستهلاك لعام ٢٠١٧/ ٢٠١٨والذى يجريه الجهاز المركزى للتعبئة العامة والاحصاء , ولقد قام الباحث بتقدير معالم نموذج الانحدار غير الخطى المتعدد فكان افضل نموذج هو النموذج اللوغاريتمى

٣- قام الباحث باستخدام برنامج MATLAB فى تدرب الشبكة العصبية الاصطناعية وذلك من خلال انشاء شبكة أمامية بتغذية مرتدة وتم تقدير المعالم والتحقق من صلاحية النموذج باستخدام أسلوب التغذية الخلفية Back propagation algorithm

بعد حساب القيم التنبؤية لاجمالى الدخل للأسر بجانب العشرة مشاهدات المستقبلية عن طريق دوال MATLAB للشبكات العصبية وكذلك برنامج SPSSلاسلوب الانحدار غير الخطى المتعدد , تم حساب دقة هذه التنبؤات بواسطة ثلاثة مقاييس إحصائية وهى (MSE, MAD, MAPE) , حيث كان المتوسط العام للأخطاء المطلقة للشبكات MAD(NN) = 31819.6 أما فى حالة اسلوب الانحدار المتعدد فكانت أقل MAD(Reg)= 17960.31, كذلك كان المتوسط العام لمربع الاخطاء للشبكات MSE(NN)= 1357793439 بينما كانت اقل منها فى حالة اسلوب الانحدار المتعدد MAD(Reg)= 1309636044 , كما كانت قيم مقياس متوسط نسبة الأخطاء المطلقة للشبكات MAPE(NN)= 46.7% , بينما كانت قيمتها اقل فى حالة اسلوب الانحدار المتعدد التقايدى MAPE(Reg)= 18.138% مما يعنى افضلية اسلوب الشبكات العصبية الاصطناعية عن الانحدار الحصين للتنبؤ بمعدلات وكانت نتيجة المقارنة هى افضلية اسلوب الانحدار غير الخطى المتعدد عن اسلوب الشبكات العصبية الاصطناعية للتنبؤ بقيمة الدخل للأسرة.

ثانياً: التوصيات :

يوصي الباحث بضرورة دراسة المقارنة بين طريقة الذكاء الاصطناعى والطرق التقليدية فى حالة العينات الصغيرة

(١-١١) المراجع :

أولا :المراجع العربية

١) **الجمال, زكريا يحيى وصابر, عمر وفيق, (٢٠١٢),** مقارنة التنبؤ باستخدام شبكة الانحدار العصبية المعممة باستخدام الشبكات العصبية وتحليل الانحدار, المجلة العراقية للعلوم الإحصائية.

٢)**المراد، عائدة يونس محمد (٢٠١٢),** قارنة بين الانحدار الكلاسيكي والشبكات العصبية الاصطناعية في التنبؤ بمستويات نتائج بحوث طلبة كلية التربية الرياضية، المجلة العراقية للعلوم الإحصائية.

٣)**دربال, أمينة, (٢٠١٤),** محاولة التنبؤ بمؤشرات الأسواق المالية العربية باستعمال النماذج القياسية دراسة حالة, مؤشر سوق دبي المالي, رسالة دكتوراه في العلوم الاقتصادية تخصص, نقود، بنوك ومالية, وزارة التعليم العالي والبحث العلمي (الجزائر).

٤)**رشاد، ندوى خزعل وزكي، عزة حازم (٢٠١٠)**، استخدام الشبكات العصبية الاصطناعية في تحليل الانحدار الخطي المتعدد، مجلة تنمية الرافدين، المجلد ٣٢، العدد ٩٩

٥)**مستور، آدم بريمه سليمان والخضر، أمل السر(٢٠١٤)**، معالجة مشكلة التداخل الخطي باستخدام تحليل المكونات الرئيسية (بالتطبيق على استهلاك الوقود بالسيارات في العام ١٩٧٥م)، جامعة السودان للعلوم والتكنولوجيا، كلية العلوم، الخرطوم، السودان.

٦)**ميشيل نجينفيتسكى ,(٢٠٠٤ م)** .

" الذكاء الصناعى دليل النظم الذكية " تعريب سرور على إبراهيم سرور , دار المريخ للنشر , الرياض , المملكة العربية السعودية, ص ٢٥٢

٧)**بحث الدخل والانفاق والاستهلاك لعام ٢٠١٨/٢٠١٧ يونيو ٢٠١٩ ,**
الجهاز المركزى للتعبئة العامة والاحصاء , المجلد الاول .

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.(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 .

3) Cybenko, g.(1989).

"Approximation by superposition of sigmoidal functions ", Mathematics of control, Signals and Systems, 2, 303-314 .

4) Demuth H. & Beale M. (1998) .

" Neural Networks Toolbox :Users Guide, Version 3.0",The Math Works inc Natick, MA .

5) 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.

6) John McDonald (1981)

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

7) Joop de Beer (1991)

"From birth expectations to birth forecasts: A partial-adjustment approach", Mathematical Population Studies: An International Journal of Mathematical Demography.

8) 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.

9) Turban, E., Aronson J. and Liang T. (2005).

"Decision Support Systems and Intelligent Systems ". 7th edition Prentice Hall .

10) Zurada ,J. M. (1992) .

"Artificial Neural Systems".1st ed. St. Paul, MN., West Publishing Company

11) www.Matlab.com

ملحق (A)

برنامج لتدريب الشبكة العصبية والتنبؤ بدخل الاسر الاجمالى باستخدام حزم MATLAB .

```
%برنامج لتدريب الشبكة العصبية
والتنبؤ بالدخل الاجمالى للاسر باستخدام
حزم MATLAB .

x=xlsread('xdata');
y=xlsread('ydata');
n=size(y'); h= 3 ;
 % انشاء وتدريب الشبكة . .....
        net =  fitnet(5,'trainlm');
         %
net.trainParam.showWindow =0 ;

[net,tr]= train(net, x,y);
% التنبؤ بقيم الدخل . .....
          y0=sim(net ,x);
```