استخدام الشبكات العصبونية لترميم قواعد البيانات عبر تقدير المدخلات اعتماداً على قيم المخرجات بأقل نسبة خطأ ممكنة

*دفاطمة شلاف

الملخص

قمنا في هذا البحث بتقدير القيم المفقودة أو القيم الارتيابية لمتغير مستقل (دخل الشبكة العصبونية) أو أكثر من متغيرات الدراسة وبالتالي ترميم قاعدة البيانات، من خلال تدريب شبكة عصبونية أمامية التغذية وباستخدام خوارزمية Levenberg-Marquardt والتي تعد تطويراً لخوارزمية الانتشار العكسي التقليدية المستخدمة في تدريب الشبكات العصبونية الأمامية، حيث قمنا بحذف بعض القيم من متغير مستقل أو أكثر ومن ثم التنبؤ بها من خلال الخوارزمية المقترحة في هذا البحث وذلك اعتماداً على القيمة الدنيا والعليا للمتغير المستقل المراد التنبؤ بقيمه المفقودة حيث تم تقسيم مجال المتغير الى نقاط قطع تصل الى(n) نقطة أو أكثر حسب الرغبة بدقة النتائج (يمكننا تقسيم المجال الى 1000 وأكثر من نقاط القطع) ومن ثم تعتبر هذه القيم مداخل للشبكة العصبونية بعد تدريبها مع بقية المتغيرات المستقلة التي لا تحوي قيماً مفقودة عند هذه المشاهدة هذه النقاط ويتم الحصول على خرج مقابل لكل نقطة قطع وبعد حساب القيمة الدنيا للفرق بين القيمة الحقيقية للهدف وخرج الشبكة تم اختيار نقطة القطع التي تعطي اقل قيمة لهذا الفرق وتكون هي تقديراً ذو دقة عالية للقيمة المفقودة. وقد وصلت دقة التنبؤ الى 0.99.

الكلمات المفتاحية: الشبكات العصبونية، خوارزمية الانتشار العكسي، خوارزمية ليفينبيرغ، القيم المفقودة، المتغير المستقل.

*مدرس- قسم الإحصاء الرياضي- كلية العلوم -جامعة حمص.

Using Neural Networks to Restore Databases by Estimating Inputs Based on Output Values with the Least Possible Error

Fatima Shallaf*

Abstract

In this research, we predict missing values or uncertainty values for one or more independent variables of the study and thus restore the database, by training a feed-forward neural network and using the Levenberg-Marquardt algorithm, which is a development of the traditional backpropagation algorithm used in training forward neural networks. By deleting some values from one or more independent variables and then predicting them through the proposed algorithm in this research, depending on the lower and upper values of the independent variable whose missing values are to be predicted, where the variable field was divided into cut-off points up to (n) points or more as desired accurately The results (we can divide the field into 1000 and more than the cut-off points) and then these values are considered as inputs to the neural network after training it with the rest of the independent variables that do not contain missing values when viewing these points. A corresponding output is obtained for each cut-off point and after calculating the minimum value of the difference between The true value of the target and the output of the network. The cut-off point that gives the lowest value for this difference is chosen, and it is a highly accurate estimate of the missing value. The prediction accuracy reached 0.99

Keywords: Neural networks, backpropagation algorithm, Levenberg algorithm, missing values, independent variable.

*Teacher, Dept. of Mathematical Statistics, Faculty of Science, Homs University.

.

1) مقدمة:

يبقى المتغير التابع هو المتغير الذي يهتم فيه الإحصـائيون للتنبؤ بقيمه ومن هذا الاهتمام نشـأت نماذج الانحدار سـواء كانت بسـيطة أو متعددة خطية أو غير خطية وجميعها معتمد على العلاقة التابعية (الارتباطية) بين المتغير التابع والمتغيرات المستقلة، أما التنبؤ بقيم المتغير المستقل بقيت حتى الآن مستقلة معتمدة فقط على قيم المتغير المسـتقل نفسـه فبنيت نماذج انحدار ذاتي وغيرها من النماذج التي تجعل هذا المتغير مستقل يعتمد على نفسه فقط،

حتى قمنا بهذا البحث الذي يعتمد على تطوير اسـتعمال الشـبكة العصـبونية لتُسـتعمل من خلال الخوارزمية المقترحة على التنبؤ بقيم المتغيرات المسـتقلة المفقودة، وذلك لأن سلامة البيانات وترميم القيم المفقودة التي تظهر في البيانات مهمة ضرورية لسلامة النتائج والاستنتاجات في تحليل المعطيات.

2) أهمية البحث:

اهتمت الأبحاث بنمذجة البيانات المدروسة من أجل استخدام النماذج للتنبؤ بالقيم المستقبلية أو ترميم قيم المتغير التابع وذلك باستخدام نماذج الانحدار التقليدية أو باستخدام الشبكات العصبونية ولكن لم يتم استخدام أي من هذه النماذج للتنبؤ بقيم المتغير المسـتقل ليعطي دقة عالية في التنبؤ بقيمه، لذا اعتمد البحث على اقتراح خوارزمية جديدة تعتمد على تدريب شبكة العصبونية (بناء النموذج التنبؤي)، ومن ثم تقدير القيم المفقودة وبالتالي جعل هذه الطريقة أحد الأدوات المساعدة في تقدير المعطيات المفقودة.

الهدف من البحث:

1- التعرّف على أهم خوارزميات تدريب الشـبكات العصـبونية لإيجاد العلاقة المعقدة بين المداخل والمخارج والتي لا يمكن نمذجتها بأساليب الانحدار.

2- اقتراح أسـلوب جديد لتقدير القيم المفقودة للمتغيرات المسـتقلة (مداخل الشـبكة العصبونية).

3- قياس دقة التقدير للقيم المفقودة بالأسلوب المقترح.

3) الشبكات العصبونية الاصطناعية Artificial Neural Networks: [1][2][3]

تُعد الشبكات العصبونية الاصطناعية تمثيلًا لطريقة عمل الجهاز العصبي في الكائنات الحية. حيث تم اكتشاف قابليتها للتعلم والتذكر والقدرة على التمييز واتخاذ القرارات من قبل العلماء. تتألف هذه الشبكات من مجموعة من الوحدات المعالجة تُسمى بالعصبونات، حيث تَمر الإشارات بين العصبونات من خلال خطوط الاتصال، ويتم تعيين وزن معين (قيمة عددية) لكل خط اتصال. وتتفاعل هذه الأوزان مع الإشارات الواردة إلى العصبون.

وتتضمن مكونات الشبكة العصبونية الاصطناعية مجموعة من العصبونات التي تكون متصلة ببعضها عبر روابط، ويتم تعيين قيم وزن لكل رابط وهذه القيم تؤثر على كيفية استجابة العصبونات للإشارات الواردة. ويتم تطبيق تابع تفعيل على كل عصبون لتحديد الإشارة الناتجة عنه. تختلف توابع التفعيل والطرق المستخدمة في الشبكات العصبونية الاصطناعية وفقًا للغرض والتطبيق.

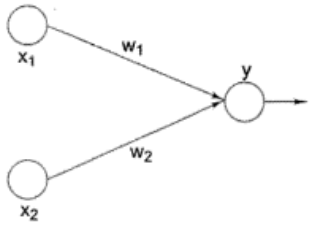
باستخدام هذه البنية، يمكن للشبكات العصبونية الاصطناعية تعلم تمثيلات معقدة للبيانات واستخدامها لمهام مثل التصنيف والتنبؤ والتشخيص. تُعتبر هذه الشبكات مكونًا أساسيًا في مجالات متعددة مثل التعلم العميق ومعالجة اللغة الطبيعية والرؤية الحاسوبية والذكاء الاصطناعي بشكل عام.

وبالتالي ستكون الشبكة العصبونية الاصطناعية مؤلفة من الأقسام التالية[4]:

1- طبقة الدخل input layer: هي طبقة الخلايا التي تستقبل إشارات الدخل من الوسط الخارجي

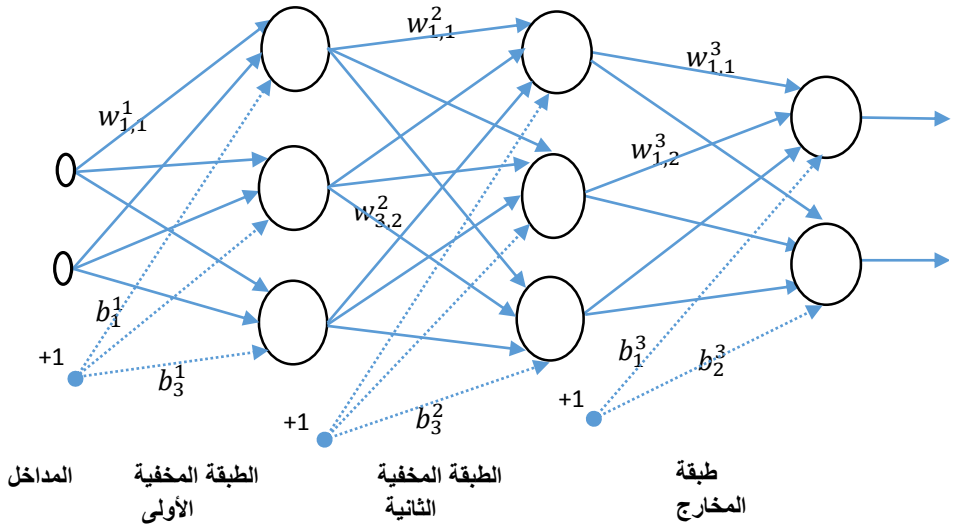
2- طبقة الخرج output layer: هي طبقة الخلايا التي تعطي القرار النهائي إلى الوسط الخارجي.

3- الطبقة المخفية hidden layer: هي طبقة (طبقات) الخلايا التي تقع بين طبقتي الدخل والخرج وتقوم بعملية المعالجة، علما أنها تكون غير موجودة في الشـــبكة العصـــبونية البســـيطة التي تكون مؤلفة من طبقة فقط (المدخلات وطبقة الخرج) كما في الشكل التالي



الشكل (1) شبكة عصبونية بسيطة وحيدة الطبقة.

أما في الشبكات العصبونية متعددة الطبقات يكون هناك طبقة مخفية واحدة أو أكثر من طبقة مخفية كما في الشكل التالي:



الشكل(2) شبكة عصبونية متعددة الطبقات.

ويتم تحديد عدد المداخل وعدد عصـــبونات طبقة الخرج من المعطيات التي يتم تحليلها لذلك إذا كان ضـــمن المعطيات أربعة متغيرات لاســـتخدامها كمدخلات،

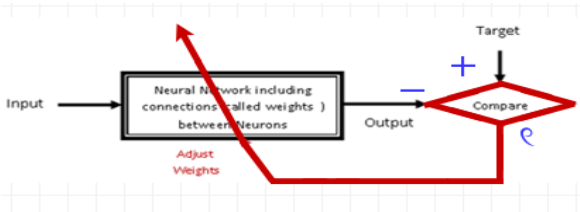
ستبنى الشبكة العصبونية بأربعة مداخل، وبالمثل إذا كان هناك سبعة متغيرات يجب دراستها كخرج للشبكة العصبونية تبنى الشبكة العصبونية بسبعة عصبونات في طبقة الخرج. أما عدد عصبونات الطبقة المخفية لم تحدد من قبل أي عالم[5] فكلها تبنى بالتجريب، فلو كان عدد عصبونات الطبقة المخفية قليلاً فإن الشبكة العصبونية لن تتعلم، وان كان عدد عصبونات الطبقة المخفية كبيراً عندها ببساطة تميل الشبكة للحفظ، لذلك يتم تحديد عدد عصبونات الطبقة المخفية بين القلة التي تجعل الشبكة لا تتعلم وبين الكثرة التي تجعل الشبكة تحفظ معطيات الدخل.[5]

4) أنواع الشبكات العصبونية حسب نوع التدريب:[1][2]

حيث تعتبر عملية التدريب من العمليات الهامة جدا بالنسبة للشبكات العصبونية، لأنها عبارة عن عملية تستطيع الشبكة العصبونية من خلالها أن تتكيف مع مدخلات محددة من أجل الحصول على استجابات مطلوبة ويمكن تصنيف أساليب التدريب المختلفة للشبكات العصبونية إلى:

5-1- التدريب المشرف عليه (المراقب) Supervised Training:[1][2]

في التعليم المشرف عليه تعطى مجموعة من بيانات التدريب (data base) المكونة من أزواج من(دخل – خرج) أي أن هذا التدريب يحتاج إلى معلم أو مشرف يتمثل بشعاع الهدف ويتم التنبؤ بالخرج لدخل جديد وبعد إدخال معطيات التدريب تعالج الشبكة المدخلات وتقارن النتائج مع النتائج المرغوبة(الخرج المرغوب) وتحسب الخطأ (الفرق بين الخرج الفعلي الذي تم الحصول عليه من المدخلات وفق خوارزمية التدريب والخرج المرغوب) وتعاد (تنشر – تبث) الأخطاء من خلال الشبكة مما يؤدي لتحديث الأوزان التي تتحكم بالشبكة، تحدث هذه العملية عدة مرات طالما أن الأوزان تتعدل.



الشكل(3) طريقة التدريب بمعلم

2-5- التدريب غير المشرف عليه (غير المراقب) Unsupervised Training:[5]

لا تحتاج الشبكات التي تستخدم هذا الأسلوب من التدريب إلى معلم أي إلى شعاع الهدف (المخرجات المرغوبة) فهي قادرة على التعلم والتصنيف للمدخلات المقدمة لها بشكل ذاتي أي تقدم المدخلات للشبكة بدون المخرجات المرغوبة، فهي نوع من أنواع شبكات التنظيم الذاتي، حيث تبني الشبكة العصبونية الاصطناعية أساليب التعليم على أساس قدرتها على اكتشاف الصفات المميزة لما يعرض عليها من أشكالٍ وأنساق وقدرتها على تطوير تمثيل داخلي لهذه الأشكال وذلك دون معرفة مسبقة وبدون عرض أمثلة لما يجب عليها أن تُنتجه وذلك على عكس المبدأ المتبع في أسلوب التعليم بوساطة معلم، ويوجد عدة أنواع للشبكات ذات خوارزميات التعلم بدون مشرف.

6- الشبكات العصبونية بتغذية أمامية feed forward networks:[4]

تنتشر هنا المدخلات من وحدات الدخل الى الطبقات المخفية فطبقة الخرج عن طريق وصلات مقترنة بأوزان كما يلي:

6-1- **المدخلات inputs:** تستقبل الاشارات الواردة من المحيط الخارجي وتمثل بالشعاع $(x_1, x_2, \dots, x_n)$ حيث يتم نقلها إلى جسم الخلية بنسب مختلفة.

6-2- **الأوزان weights:** قيم عددية قابلة للزيادة والنقصان وفق معادلات رياضية محددة ويرمز لها ب$(w_1, w_2, \dots, w_n)$، حيث يتم اعطاؤها قيم ابتدائية في الخطوة الأولى للتدريب ثم يتم تعديل هذه القيم حتى الوصول الى أصغر قيمة ممكنة للخطأ الذي يتمثل بالفرق بين الخرج الفعلي للشبكة العصبونية والخرج المرغوب.

6-3- **تابع التجميع summation function:** يقوم بجمع المدخلات الموزونة weighted function وفق العلاقة:

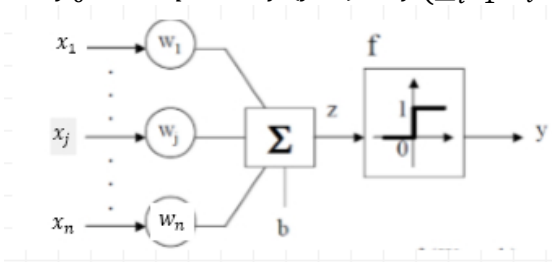
$$y_{in} = \sum_{i=1}^{i=n} w_i x_i = w_1 x_1 + w_2 x_2 + \cdots + w_n x_n \mp b \qquad (1)$$

6-4- **تابع التفعيل (التنشيط)activation function:** يقوم بتحويل حاصل جمع المدخلات الموزونة إلى إشارة رياضية محددة حسب طبيعة المشكلة لينتج الخرج النهائي للخلية العصبية كما يلي:

$$y_{in} = \sum_{i=1}^{n} w_i x_i \mp b \qquad (2)$$

فيكون خرج الشبكة العصبونية بتطبيق تابع التفعيل على y_{in} كما في العلاقة التالية:

$$\hat{y}_0 = output = f(yin) = f\left(\sum_{i=1}^{i=n} w_i x_i\right) \qquad (3)$$



الشكل (4) الشكل التفصيلي للشبكة العصبونية perceptron Network.

7- خوارزميات تدريب الشبكات العصبونية:

لدينا عدة خوارزميات تدريب للشبكات العصبونية نذكر منها: [5][4]

1- شبكة التكيف الخطي ADALINE NN : خوارزمية التدريب : delta

2- شبكة التكيف الخطي متعددة الطبقات MADALINE NN : هي عبارة عن تركيب لشبكة التكيف الخطي الـ Adaline . ولهذه الشبكة خوارزميتي تدريب MRI,MRII

3- خوارزمية الانتشار العكسي Back Propagation Algorithm.

8- خوارزمية الانتشـــار العكســـي Back Propagation Algorithm [5][6]

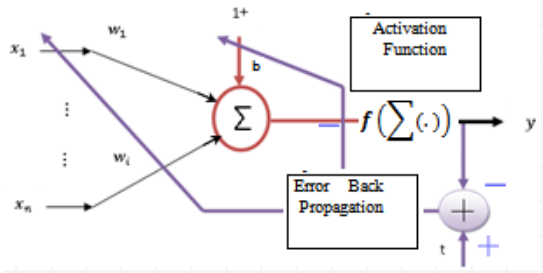
تم اقتراح هذه الخوارزمية من قبل Paul Werbos في الســـبعينات، ولكن لم تســـتعمـل إلا لعـام 1986م من قبـل Rumelhart و McClelland وهي من أهم خوارزميات تدريب شـبكة عصـبونية تعتمد على الانتشـار العكسـي للخطأ من طبقة

الخرج باتجاه الطبقات السابقة (المخفية إلى الدخل)، تُستخدم لتعليم الشبكات العصبونية ذات التعلم بإشراف، والجدير بالذكر أن تابع التفعيل المستخدم في هذه الخوارزمية هو تابع الـSigmoid، كما تعتبر خوارزمية الانتشار العكسي من الخوارزميات السريعة لكنها تأخذ وقت في إيجاد مجموعة الحلول المثلى للأوزان، ولا يمكن استخدامها إلا بوجود طبقة أو اكثر من الطبقات المخفية.

تطوير فكرة الانتشار الخلفي لتدريب الشبكات العصبونية متعددة الطبقات لعبت الدور الأساس في إبراز الشبكات العصبونية الصناعية كأداة لحل الكثير من المشاكل على نطاق واسع.

وتهدف عملية تدريب هذه الشبكات الوصول إلى حالة من التوازن بين قابلية الشبكة على الاستجابة لعينات الدخل التي تستخدم في عملية التدريب وقدرتها على إعطاء استجابة جيدة لدخل مشابه لكن غير مطابق لذلك الدخل المستخدم في التدريب.

يتضمن تدريب الشبكة بطريقة الانتشار الخلفي ثلاثة مراحل: مرحلة الانتشار الامامي والانتشار الخلفي ثم مرحلة تعديل الأوزان.وبعد مرحلة التدريب تبدأ مرحلة اختبار الشبكة والتي تتضمن مرحلة واحدة فقط هي طور الانتشار الأمامي.



الشكل (5): تدريب الشبكات الأمامية متعددة الطبقات

9- تطوير خوارزمية الانتشار العكسي: [7][6][4]

ان لخوارزمية الانتشار العكسي العديد من التطويرات والتطور الاهم هو خوارزمية Levenberg-Marquadt، والتي تقلل من مجموع مربعات الخطأ $t_k - y_k$ (الفرق بين الخرج المطلوب و الخرج الفعلي)[7] وهي تحقق نسبة تقارب أفضل وتتطلب ذاكرة أقل وأسرع من غيرها من خوارزميات تدريب الشبكات العصبونية الأخرى مثل خوارزمية

الانتشار العكسي (Error Back propagation Algorithm) للخطأ حيث أن هذه الخوارزمية تكون غالباً بطيئة في المشاكل العملية حيث أن خوارزمية Levenberg-Marquardt يمكن أن تكون أسرع منها من 10 حتى 100 مرة وهي تعتبر تطوير لخوارزمية الانتشار العكسي للخطأ [6][4]

10- خوارزمية Levenberg- Marquardt:[7]

وهي من طرائق ايجاد الحلول المثلى لأوزان الشبكة العصبونية من خلال ايجاد مجموع المربعات الصغرى المعرفة بالشكل:

$$E(W_j) = \frac{1}{2}\sum_i^N \sum_{k=1}^K (t_{ik} - y_{ik})^2 \qquad (6)$$

علماً أن K: عدد عصبونات طبقة الخرج، N: عدد أنماط التدريب (iteration) .

j: تعبر عن كل مسح لكامل البيانات (epoch)

وبترميز للخطأ بالشكل التالي:

$$e_i(W_j) = \sum_{k=1}^K (t_{ik} - y_{ik}) \qquad (7)$$

وبالتالي يكتب تابع الكلفة بالشكل التالي:

$$E(W_j) = \frac{1}{2}\sum_i^N e_i^2 \qquad (8)$$

وحسب طريقة نيوتن تكون الحلول المتتالية معرفة وفق العلاقة التالية:

$$W_{j+1} = W_j - H_j^{-1} g_j \qquad (9)$$

حيث أن: $H_j = \nabla^2 E(W)|_{W=W_j}$ وعندما نفترض أن $E(W_j)$ هو تابع مجموع المربعات:

$$E(W_j) = \frac{1}{2}\sum_i^N {e_i}^2 = e'e \qquad (10)$$

عندها يكون العنصر ذو الترتيب j للمشتق معرف كما يلي:

$$[\nabla E(w)]_j = \frac{\partial E(W_j)}{\partial W_j} = 2\sum_{i=1}^N e_i \frac{\partial e_i(X)}{\partial W_j} \qquad (11)$$

وبالتالي يكون شكل مصفوفة المشتق كما يلي:

$$\nabla E(w) = 2J^T(W)e(W) \qquad (12)$$

حيث أن:

$$J(W)=\begin{bmatrix}\frac{\partial e_1(W)}{\partial W_1} & \frac{\partial e_1(W)}{\partial W_2}\cdots & \frac{\partial e_1(W)}{\partial W_n}\\ \frac{\partial e_2(W)}{\partial W_1} & \frac{\partial e_2(W)}{\partial W_2}\cdots & \frac{\partial e_2(W)}{\partial W_n}\\ \vdots & \vdots & \vdots\\ \frac{\partial e_N(W)}{\partial W_1} & \frac{\partial e_N(W)}{\partial W_2}\cdots & \frac{\partial e_N(W)}{\partial W_n}\end{bmatrix} \quad (13)$$

n : عدد الأوزان ضمن الشبكة العصبونية.

حيث تدعى $J(W)$ بمصفوفة جاكوبي.

ويتم كتابة مصفوفة هيسيان بطريقة تقريبية وفق العلاقة التالية:

$$\nabla^2 E(W)\cong 2J^T(W)J(W) \quad (14)$$

وبالتالي حسب طريقة Gauss_Newton يمكن كتابة مايلي:

$$W_{j+1}=W_j-\left[2J^T(W_j)J(W_j)\right]^{-1}2J^T(W_j)e(W_j) \quad (15)$$

$$=W_j-\left[J^T(W_j)J(W_j)\right]^{-1}J^T(W_j)e(W_j) \quad (16)$$

ونلاحظ أن طريقة نيوتن القياسية ليست بحاجة الى حساب المشتق الثاني. وأحد مشاكل Gauss_Newton أن مصفوفة Hessian $H=J^TJ$ قد لاتملك مقلوب لذلك نستخدم مصفوفة تقريبية لHessian معرفة بالشكل التالي:

$$G=H+\mu I \quad (16)$$

μ: برامتر ثابت يضاف في حال كانت H تملك محدداً يساوي الصفر وهو رقم صغير جداً لا يؤثر على قيمة العناصر الموجودة في المصفوفة H

وبالتالي تعطى الأوزان لهذه الخوارزمية كما يلي:

$$W_{j+1}=W_j-\left[J^T(W_j)J(W_j)+\mu I\right]^{-1}J^T(W_j)e(W_j) \quad (17)$$

$$\Delta W_k=-\left[J^T(W_j)J(W_j)+\mu I\right]^{-1}J^T(W_j)e(W_j) \quad (18)$$

حيث I هي المصفوفة الواحدية لها نفس درجة المصفوفة H،J هي مصفوفة Jacobian لأخطاء الخرج m المتعلقة بالأوزان n للشبكة العصبونية.

11- الخوارزمية المقترحة من أجل تقدير القيم المفقودة والشاذة:

إذا كان لدينا X_{ij} متحولات مستقلة (K متحول مستقيل يتألف كلاً منها من nقيمة)

$$i=1,2,\dots,n$$
$$j=1,2,\dots,k$$

وكان y_i متحول تابع:

فإن الخوارزمية المقترحة توصف بالخطوات التالية:

1- حذف جميع الأسطر من البيانات التي توجد فيها قيم مفقودة في أحد المتحولات المستقلة.

2- بناء شبكة عصبونية متوافقة مع البيانات المتبقية بعد عملية الحذف بجعل المداخل هي X_{ij} والخرج هو y_i .

3- تدرب الشبكة العصبونية والحصول على شبكة مدربة واستخلاص قيم الأوزان.

4- ترميم قاعدة المعطيات قيمة تلو أخرى فإذا فرضنا أن المتغير رقم r ونريد التنبؤ بالقيمة رقم i تكون المعالجة كما يلي:

$$n_1 = \min(X_{ir})$$
$$n_2 = \max(X_{ir})$$

تقسيم المجال $[n_1 - n_2]$ إلى عدد من القيم mبخطوة ثابتة مقدارها α وكلما كانت m كبيرة كانت العملية أدق وأفضل وأكثر شفافية.

نثبت باقي قيم المتحولات الأخرى على وضعها في سطر المقابل للقيمة التي يتم التنبؤ بها في المتغير r ونحصل على مصفوفة جديدة نتيجة التقسيم كما يلي:

$$\begin{bmatrix} X_{i1} & \dots & n_1 & X_{ik} \\ X_{i1} & \dots & n_1 + \alpha & X_{ik} \\ X_{i1} & \dots & n_1 + 2\alpha & X_{ik} \\ \vdots & \vdots & \vdots & \vdots \\ X_{i1} & \dots & n_2 & X_{ik} \end{bmatrix}$$

نعوض المصفوفة في الشبكة العصبونية المدربة:

نحسب: $Z_i = \hat{y}_i - y_i, i = 1,2, \dots, m$

التقدير الأمثل لـ X_{ir} هي لقيمة التي تقابل $\min(Z_i)$

التطبيق العملي:

تناولت الدراسة بيانات الطقس في خمس محطات في مدينة حلب (محطة الشيخ سعيد جنوب حلب_ محطة طريق المسلمية شمال حلب_ محطة مركز المدينة وسط حلب _ محطة حلب الجديدة غرب حلب_ محطة المطار المدني شرق حلب) ابتداءً من شهر كانون الأول من عام 2009 حتى شهر أيار من عام 2011 حيث اعتمدنا الحرارة الداخلية متغيراً تابعاً واعتمدنا أربع متغيرات مستقلة تؤثر على الحرارة الداخلية، وهي درجة الحرارة وسرعة الرياح والرطوبة الجوية

يمكن تعريف متغيرات الدراسة بالجدول التالي:

الجدول (1.3) متغيرات الدراسة

المتغير	الرمز	
التابع	y	**الحرارة الداخلية**
المستقل	X_1	**درجة الحرارة**
المستقل	X_2	**سرعة الرياح**
المستقل	X_3	**اتجاه الرياح**
المستقل	X_4	**الرطوبة الجوية**

بحذف بعض القيم الموجودة في المتغير المستقل الاول وتطبيق الخوارزمية المقترحة تم الحصول على النتائج التالية:

رقم القيمة	$X1(69)$	$X1(100)$	$X1(116)$
القيمة الحقيقية	23.962	11.377	17.167
القيمة المقدرة	24.10501	11.25668	17.20229
دقة التقدير	0.994067	0.989424	0.997177

النتائج:

تم ترميم قاعدة البيانات من خلال تقدير القيم المفقودة وذلك بتدريب الشبكة العصبونية وتطبيق الخوارزمية المقترحة.

حصلنا على دقة عالية في تقدير القيم المفقودة تصل الى احتمال الحدث الأكيد.

التوصيات:

1- الاعتماد على الشـــبكات العصـــبونية في معالجة القيم المفقودة والمشـــاكل الإحصائية الأخرى.

2- اعتماد طريقة الشــبكات العصــبونية في بناء النماذج الرياضــية واســتخدامها لأغراض التنبؤ.

المراجع:

[1]− Michael Negnevitsky, 2005- **Artificial Intelligence A Guide to Intelligent Systems**- 2ed **Edition.Addison-Wesley p435 .**

[2] Zurada M. Jacek (1995) , **"Introduction to Artificial Neural Systems**" , PWS.

[3]−Ben Coppin,2004-**Artificial Intelligence Illuminated**- Jones And Bartlett Publishers

[4] Sivanandam S N, Sumathi S, Deepa S N.,2006-**introduction to Neural Networks Using Matlab 6.0**-Tata Mcgraw_Hill Publishing Company Limited New Delhi.p549.

[5] فاطمة**شـــلاف** 2019 **معالجة المعطيات المفقودة والشـــاذة من خلال تطوير خوارزمية شبكة ذكاء صنعي متقدمة_ جامعة حلب، سوريا، رسالة دكتوراه.**
Department of Information Technology Computer Systems

[6] B. Schiele and A. Waibel**. " Gaze tracking based on face-color"**. Presented in International Workshop on Face and Gesture Recogntion, Zurich, July 2001

[7]- Martin T. Hagan. Hagan, Howard B. Demuth, Mark Hudson Beale, Orlando De Jesús.,1996-**Neural Network Design** 2ed **Edition**. Martin T. Hagan.