دار المنظومة
DAR ALMANDUMAH
الرواد في قواعد المعلومات العربية

العنوان:	الشبكات العصبية و تطبيقات ادارة الاعمال
المصدر:	المجلة العلمية للاقتصاد والتجارة
الناشر:	جامعة عين شمس - كلية التجارة
المؤلف الرئيسي:	عبدالعال، مدحت محمد أحمد
المجلد/العدد:	ع1
محكمة:	نعم
التاريخ الميلادي:	2004
الشهر:	يناير
الصفحات:	465 - 494
رقم MD:	110730
نوع المحتوى:	بحوث ومقالات
قواعد المعلومات:	EcoLink
مواضيع:	تكنولوجيا المعلومات ، الشبكات العصبية الصناعية ، ادارة الاعمال ، التحليل الاحصائي، الذكاء الصناعي، تكنولوجيا الاتصالات ، التعلم الذاتي
رابط:	http://search.mandumah.com/Record/110730

© 2020 دار المنظومة. جميع الحقوق محفوظة.
هذه المادة متاحة بناء على الإتفاق الموقع مع أصحاب حقوق النشر، علما أن جميع حقوق النشر محفوظة. يمكنك تحميل أو طباعة هذه المادة للاستخدام الشخصي فقط، ويمنع النسخ أو التحويل أو النشر عبر أي وسيلة (مثل مواقع الانترنت أو البريد الالكتروني) دون تصريح خطي من أصحاب حقوق النشر أو دار المنظومة.

الشبكات العصبيه وتطبيقات ادارة الاعمال

الدكتور مدحت محمد احمد عبد العال

٤٦٥

استاذ الاحصاء والرياضه المساعد

قسم الاحصاء والرياضه

كلية التجاره جامعة عين شمس

ملخص البحث:

أن الشبكات العصبيه تم ابتكارها اصلاً من علم النيورونات لتستخدم كاداه قويه للتحليل الاحصائى. ونجد أن اهم ميزه فى الشبكات العصبيه هو قدرتها على التعلم اعتماداً على عدد محدد من المشاهدات. أن المراجع العلميه الخاصه بالشبكات العصبيه تعتبر المصطلح "تدريب" انه معلومات يمكن اشتقاقها من العينه وتعميمها لمشاهدات اخرى غير مرئيه. ونجد ان الشبكات العصبيه يطلق عليها اسم "الة التعلم". ويمكن اعتبار الشبكات العصبيه كعميل يتعلم من البيئه وبالتالى يمكنه مواجهة اى سلوك ما اعتماداً على ما تعلمه من عدد محدد من المشاهدات. ومن هذا المنطلق فأن الباحث لن يتطرق الى المفهوم البيولوجى للشبكات العصبيه ولكن سوف يقتصر هذا البحث على المفهوم الرياضى والتطبيقات الاحصائيه للشبكات العصبيه.

1. المقدمه:

تعتبر الشبكات العصبيه الصناعيه ANN احد طرق الذكاء الصناعى AT والتى تم تصميمها طبقا لتصميم المخ البشرى. ان بداية علم الشبكات العصبيه كان يسير فى اتجاهين مختلفين:

الاتجاه الاول وهو اتجاه بيولوجى حيث تم اكتشاف واستخدام الشبكات العصبيه لعمل محاكاه بسيطه للمخ البشرى وذلك لاختبار وظائفه، أما الاتجاه الثانى فكان استخدام الشبكات العصبيه كنظام تقنيه لتشغيل او تحليل المعلومات المعقده (Zahedi, 1993).

ان السبب وراء تفضيل الشبكات العصبيه على طرق التحليل الاحصائى التقليديه هو قدرة الشبكات العصبيه على تحليل البيانات المشوشه أو التى تحتوى على قيم متطرفه كثيره وايضا امكانية الوصول الى حل محدد للمشاكل التى ليس لها حل محدد وايضا امكانية التحليل والتعلم من البيانات التاريخيه (السلاسل الزمنيه) وحيث ان هذه المميزات تتوافر عند استخدام الشبكات العصبيه فانها قد اظهرت نجاح فى التنبؤ ببيانات السلاسل الزمنيه الماليه والتى كانت لها خاصية عدم الثبات والتقلبات بشكل كبير ولكن يجب ان لاننسى ان هناك عيب فى استخدام الشبكات العصبيه وهو عدم القدره على تطبيق الاختبارات الجوهريه الاحصائيه للنماذج المقترحه وللمعلمات المقدره (Refenes, Abu-Mostafa, Moody, Weigend, 1996) وبالرغم من هذا العيب فان نتائج العديد من الابحاث اوضحت ان الشبكات العصبيه يمكنها حل معظم المشاكل أن لم يكن جميع المشاكل بكفاءه اعلى من الطرق الاحصائيه التقليديه. وقد قدم Masters فى عام 1995 اثبات

رياضى يؤكد ان الشبكه العصبيه ذات الثلاث طبقات والتى لها دالة تحويل عشوائيه يمكنها التقريب لأى داله غير خطيه.

2- تاريخ الشبكات العصبيه:

لقد قام كل من McCollough and Pitts فى عام 1943 بتناول اول بحث للشبكات العصبيه. وفى عام 1958 قام Rosenblatt بتصميم نظام التعلم لنموذج يسمى عنصر الادراك Perceptron ولقد زاد الاهتمام بدراسة الشبكات العصبيه عام 1969 عنما قام كل من Minsky and Papert باصدار كتاب عن الشبكات يفيد انه يمكن استخدام الشبكات العصبيه لتقدير عدد محدد من الدوال. وبظهور هذا الكتاب وتطوير نموذج هوبيفلد Hopfield وطريقة الارتداد الخلفى backpropagation algorithm والتقدم السريع جداً فى تقنية الحسابات زاد الاهتمام باستخدام الشبكات (Ahmadian 1995).

3 – اساسيات الشبكات العصبيه:

ان الشبكات العصبيه تتكون من طبقين أو اكثر او تتكون من مجموعات من عناصر التشغيل التى تسمى نيورون neuron، ان المصطلح نيورون neuron يشير الى الوحده الاساسيه فى الشبكات العصبيه والتى يمكن استخدامها لتشغيل البيانات.

ان النيورونات تتصل بالشبكه العصبيه بحيث ان مخرجات اى نيورون تعتبر المدخلات لنيورون آخر. وهذا يعنى وجود اتصال بين النيورونات وبالتالى يوجد اتجاه لهذة الاتصالات. وطبقا لاتجاه الاتصال فان طريقة الاتصال قد تكون اتصال احادى one-directional او اتصال ثنائى bi-directional، وطبقا لكثافة الاتصال فان الاتصال قد يكون اتصال نشط أو اتصال غير

نشط. ويتم تقسيم النيورونات فى مجموعات تمثل الطبقات ونجد ان هناك ثلاث انواع من الطبقات هى طبقة المدخلات Input layer والطبقه الخفيه Hidden layer وطبقة المخرجات Output layer.

ان مهمة طبقة المدخلات هى استقبال البيانات من مصدر خارجى ثم أرسالها الى الطبقه الخفيه. داخل الطبقه الخفيه يتم تشغيل البيانات ثم ارسال نتائج هذا التشغيل الى طبقة المخرجات حيث يتم مقارنة مخرجات الشبكه بالمخرجات الحقيقية وذلك لحساب مقدار الخطأ فى التقديرات. يلى ذلك ارجاع الخطأ الذى تم حسابه مره اخرى الى الشبكه ويتم تعديل اوزان الاتصال Connection Weights بين النيرونات ثم يتم تشغيل البيانات مره اخرى واعادة الخطوات السابقه عدة مرات حتى يتم الحصول على مخرجات الشبكه قريبه جدا للمخرجات الحقيقية وبهذا يتم تصغير الخطأ والذى يمثل الفرق بين مخرجات الشبكه والمخرجات الحقيقيه.

ان اوزان الاتصال تمثل قوة الاتصال بين النيورونات فعلى سبيل المثال اذا كان النيورون j يتصل بالنيورون i فان W_{ji} تمثل وزن الاتصال بين النيورون j والنيورون i وكذلك فان W_{ij} هى وزن الاتصال بين النيورون i والنيورون j.

واذا كان النيورون i متصل بالنيورونات 1 و2 و 3.....n فان اوزان الاتصال يتم حفظها فى المتغيرات للآتيه $W_{11}, W_{21}, W_{31}, ...W_{ni}$. ونجد ان النيورون يستقبل العديد من المدخلات طبقا لعدد النيورونات المتصله به، ثم يقوم هذا النيورون باستخدام هذة المدخلات لانتاج مخرجات واحده ترسل الى نيورون آخر وذلك طبقا لداله التحويل المستخدمه (انظر الفقره (2.4.3).

ان عمليات تصميم تشغيل الشبكه العصبيه يتكون من اربع خطوات (Zahedi, 1993) هى:

1- تحديد النيورونات فى مختلف الطبقات.

2- تحديد طبيعة الاتصال بين النيورونات، ونجد أن الاتصال قد يكون اتصال بين الطبقات Inter-layer او اتصال داخل الطبقه الواحده Intera–layer (انظر الفقره 1.2.3 الفقره 2.2.3)

3- تحديد اتجاه استقبال المدخلات واتجاه انتاج المخرجات.

4- تحديد قاعدة التعلم لتعديل اوزان الاتصال.

ونجد انه لتركيب أو لبناء شبكه عصبيه فأنه يجب توافر المتطلبات التاليه:

1- فئة البيانات.

2- طريقة الاتصال بين النيورونات.

3- الاتصال بين المدخلات والمخرجات.

4- دوال المدخلات ودوال التحويل.

5- طرق التعلم.

6- معلمات التعلم.

1.3 فئة البيانات:

عند استخدام الشبكه العصبيه فان العينه يتم تقسيمها الى ثلاثه اجزاء هى فئة التدريب Training set وفئة الصلاحيه Validation set وفئة الاختبار Test set ونجد ان هناك تشابه بين فئة الصلاحيه وفئة الاختبار ولقد قام Ripley فى عام 1996 بعرض تعريف محدد لكل فئه من هذة الفئات وذلك فى صفحة 354 وهذه التعريفات هى:

1.1.3. فئة التدريب:

هى مجموعه من المشاهدات تستخدم لتقدير معلمات النموذج المقترح.

2.1.3. فئة الصلاحيه:

هى مجموعه من المشاهدات تستخدم لتعديل تقديرات المعلمات التى تم حسابها بأستخدام فئة التدريب، ويمكن استخدامها ايضا لتحديد عدد الطبقات الخفيه فى الشبكه.

3.1.3. فئة الاختبار:

هى مجموعه من المشاهدات تستخدم لتعميم النتائج أو لاختبار التقديرات التى تم حسابها باستخدام فئة التدريب وتم تعديلها باستخدام فئة الصلاحيه.

وقد اصدر Bishop فى عام 1995 كتاب يعتبر من المراجع الاساسيه لدراسة الشبكات العصبيه وقد ذكر فى كتابه ان الهدف من الشبكات العصبيه هو الحصول على افضل تمثيل للبيانات المستخدمه وان ابسط اسلوب للمقارنه بين الشبكات العصبيه المختلفه التى يمكن بنائها لتمثيل مجموعة بيانات ما هو تقدير داله الخطأ باستخدام مجموعة بيانات تكون مستقله عن البيانات التى يتم استخدامها لتدريب الشبكه، وهناك العديد من الشبكات العصبيه يمكن تدريبها من خلال تصغير او تدنيه دالة الخطأ ونجد ان المفاضله بين الشبكات العصبيه يتم من خلال مقارنة دالة الخطأ وذلك باستخدام فئة الصلاحيه وان الشبكه صاحبة اقل خطأ يتم اختيارها ونجد ان هذا الاجراء يمكن ان يؤدى الى ما يسمى بمشكلة Over fitting لفئة الصلاحيه وفى هذه الحاله يتم استخدام الفئه الثالثه من البيانات وهى فئة الاختبار للمفاضله بين الشبكات العصبيه.

وبالرجوع الى تعريف Bishop فى عام 1995 نجد ان فئة الاختبار لايمكن استخدامها للمفاضله بين الشبكات العصبيه ولكن يمكن استخدامها للحصول على تقدير غير متحيز للخطأ المعمم. وقد أكد Bishop على أن فئة الصلاحيه تستخدم للمفاضله بين الشبكات المختلفه ان تقدير الخطأ الناتج من استخدام فئة الصلاحيه يعتبر تقدير متحيز.

ونجد ان هناك بعض طرق التدريب مثل طريقة التوقف المبكر Early stopping تتطلب وجود فئة الصلاحيه والتى يتم استخدامها لتدريب الشبكه وهناك طرق اخرى مثل طريقة الجوازيه الكبرى Maximum Likelihood لا تتطلب وجود فئة الصلاحيه ولهذا فان هذه الطريقه (الجوازيه الكبرى) تقوم بدمج فئة التدريب مع فئة الصلاحيه لاستخدام طريقة التوقف المبكر ولقد اقترح كل من Cherkassky, Friedman, and Wechsler فى عام 2001 تكوين فئه تسمى فئة التصميم Design set والتى يمكن استخدامها لتعديل اوزان الشبكه.

2.3. طريقة الاتصال بين النيورونات:

ان الاتصال فى الشبكات العصبيه يمكن ان يتحقق بين طبقتين وهو مايسمى بالاتصال الخارجى Inter–Layer Connections ويمكن ايضا ان يكون داخل الطبقه الواحده وهو مايسمى بالاتصال الداخلى Intra–Layer Connections.

1.2.3. الاتصال الخارجى Inter–Layer Connections:

لقد ذكر Zahedi فى عام 1993 ان هناك عدة طرق للاتصال بين الطبقات وبعضها البعض وهذه الطرق هى:

(أ) **اتصال كامل Fully connected:**

ان كل نيورون فى الطبقه الاولى يتصل بكل نيورون فى الطبقه الثانيه.

(ب) **اتصال جزئى Partially connected:**

ان اى نيورون فى الطبقه الاولى ليس بالضروره ان يتصل بكل النيورونات فى الطبقه الثانيه.

(ت) **الاتصال الامامى Feed-forward:**

ان الاتصال بين النيورونات يكون من النوع احادى الاتصال حيث ان النيورون فى الطبقه الاولى يقوم بارسال مخرجاته الى النيورونات فى الطبقه الثانيه ولكن لايقوم باستقبال اى مخرجات من الطبقه الثانيه أو باسلوب آخر ان نقل المخرجات يتم فى اتجاه واحد وهو الاتجاه الامامى.

(ث) **الاتصال الثنائى Bi-directional:**

فى هذه الطريقه يقوم النيورون فى الطبقه الثانيه بارسال المخرجات مره اخرى الى النيورونات فى الطبقه الاولى.

(ج) **الاتصال الهرمى Hierarchical:**

فى هذه الطريقه تقوم النيورونات فى الطبقات الادنى بالاتصال فى النيورونات فى الطبقات الاعلى.

(ح) **الاتصال الرنينى (المشروط) Resonance:**

الاتصال فى هذه الطريقه يعتبر ثنائى حيث ان النيورونات تستمر فى ارسال المعلومات بين الطبقات حتى يتحقق شرط معين.

2.2.3. الاتصال الداخلى intra-layer connections:

وهو الاتصال بين النيورونات داخل الطبقه الواحده وهناك عدة طرق للاتصال الداخلى وهى:

(أ) الاتصال المتكرر Recurrent

هذا النوع من الاتصال قد يكون اتصال كلي او اتصال جزئي. ان هذا النوع من الاتصال يتحقق عندما تقوم النيورونات بنقل مخرجاتها الى النيورونات الاخرى بعد استقبال مدخلاتها من طبقه اخرى. هذا الاتصال يستمر حتى تصل الشبكه الى وضع الثبات وفى حالة تحقق ثبات الشبكه فهذا يسمح للنيورونات بارسال مخرجاتها الى الطبقه التاليه.

(ب) أتصال نشط فى المركز وغير نشط خارج المركز

On Center /Off Surround

فى هذه الطريقه نجد ان النيورون فى احد الطبقات له اتصال نشط مع النيورونات المجاوره ولكن هناك اتصال غير نشط مع النيورونات فى الطبقات الاخرى.

3.3. الاتصال بين المدخلات والمخرجات:

ان تصميم الشبكات العصبيه يمكن ان يعتمد على طبيعة الاتصال بين المدخلات والمخرجات والتى يمكن ان تكون كما يلى:

شبكه ذاتية الترابط أو الاتصال Auto- Association

فى هذه الطريقه نجد ان متجه المدخلات هو نفسه متجه المخرجات.

شبكة غير ذاتية الترابط أو الاتصال Hetero-associative

فى هذه الطريقه نجد ان متجه المخرجات يختلف عن متجه المدخلات.

ان الشبكات العصبيه من النوع Auto-Associative يمكن استخدامها فى تحديد نمط البيانات، تشغيل الاشارات، تنقية القيم المتطرفه واكتشاف نمط البيانات.

4.3. المدخلات ودوال التحويل:

ان الدوال الاساسيه التى تستخدم فى الشبكات العصبيه هى داله المدخلات ودوال المخرجات او دوال التحويل.

1.4.3. دوال المدخلات (دالة المجموع):

عندما يستقبل النيورون المدخلات من الطبقه السابقه فان قيمة هذه المدخلات يتم حسابها طبقا لداله المدخلات المستخدمه وعادة مايتم استخدام دالة المجموع Summation Function. ونجد ان دالة المجموع للنيورون i يتم تحديدها عن طريق ضرب المخرجات التى تم ارسالها من النيورون j الى النيورون i يرمز لها بالرمز $Output_i$ بأوزان الاتصال بين النيورون i والنيورون j ثم جمع نواتج الضرب وذلك طبقا للمعادله التاليه.

$$input_i = \sum_{j=1}^{n} \left(W_{ji} * output_j \right)$$

حيث ان n هى عدد النيورونات فى الطبقه والتى قامت بارسال مخرجاتها وتم استقبالها بواسطه النيورون i وبطريقه اخرى فإن $Input_i$ للنيورون i هي المجموع المرجح للمخرجات باوزان الاتصال والتي تم إرسالها لهذا النيورون.

إن قيم المدخلات يمكن أن يتم تحويلها إلي قيم معيارية باستخدام دالة التوزيع الطبيعى وهذه القيمة المعيارية قد تتراوح بين صفر وواحد صحيح أو سالب واحد وموجب واحد وذلك لتجنب تأثير القيمة المتطرفة في المدخلات ولهذا فأن تحويل المدخلات إلي قيم معيارية يعتبر إجراء مفضل لاغلب الشبكات العصبية كما ورد في بحث كل من Cross, Harrison, and Kennedy, في عام 1995.

3. 4. 2. دالة التحويل:

بعد استقبال المدخلات طبقا لدالة الجمع التي يتم استخدامها فإن مخرجات النيورون يتم حسابها ثم ترسل الي الطبقة التالية التي يتصل بها هذه النيورون. يلي ذلك حساب مخرجات النيورونات طبقا لدالة التحويل المستخدمة ونجد أن اكثر دوال التحويل استخداما هي

1 – Step Function.

2 – Sign Function.

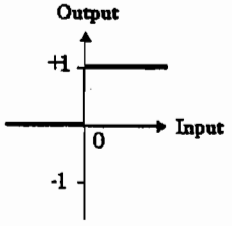
3 – Sigmoid Function.

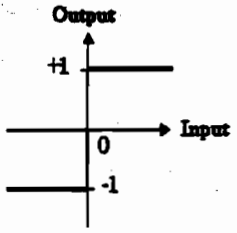
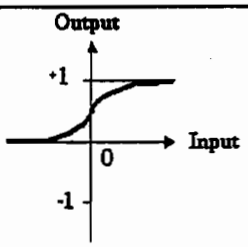
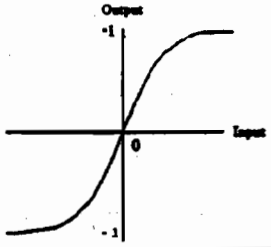
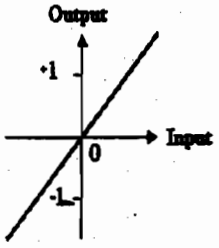
4 – Hyperbolic- Tangent Function.

5 – Linear Function.

الجدول الآتي رقم (1) يعرض المعادلات الرياضية لهذه الدوال والرسم البياني لكل دالة.

جدول رقم (1) اكثر دوال التحويل استخداماً والرسم البياني لكل داله.

The Function	The Formula	The graph
Step function	$output_i = \begin{cases} 0 & when\ input_i \le T \\ 1 & when\ input_i > T \end{cases}$	

Sign function	$output_i = \begin{cases} 1 & when\ input_i > 0 \\ 0 & when\ input_i = 0 \\ -1 & when\ input_i < 0 \end{cases}$	
Sigmoid function	$output_i = \frac{1}{1+e^{input_i}}$	
Hyperbolic tangent function	$output_i = \frac{e^{input_i} - e^{-input_i}}{e^{input_i} + e^{-input_i}}$	
Linear function	$output_i = g * input_i$	

- من هذا الجدول نجد أن مخرجات Step Function يتم حسابها طبقا للمعادلة الرياضية المذكورة حيث أن T القيمة الحدية Threshold وأن T تنتمي إلي فئة الأعداد الطبيعية.

- تعتبر دالة Sign Function صورة خاصة من دالة Step Function ولكن قيمة T تساوي صفر ولقد تم استخدام هذه الدالة في أول شبكة عصبية والتي تسمي شبكة عنصر الادراك Perceptron بواسطة Rosenblat في عام 1985.
- تعتبر دالة Sigmoid Function من اكثر دوال التحويل استخداما في الشبكات العصبية ونجد أن هذه الدالة تعطي قيم مستمرة بين الفترة صفر وواحد صحيح.
- تعتبر دالة Hyperbolic Tangent Function صورة خاصة من Sigmoid Function ونجدا أن الرسم البياني لهذه الدالة يتشابه مع الرسم البياني لدالة Sigmoid والاختلاف هو أن هذه الدالة Hyperbolic Tangent Function تتراوح بين سالب واحد وموجب واحد.
- إن الدالة الخطية تمر خلال نقطة الأصل وتعتمد علي قيمة المعلمه g حيث أن قيمة هذة المعلمه يساوى مقلوب القيمة الحديه T.

3. 5 – طرق التعلم:

إن تعلم الشبكة أو تدريب الشبكة هو عملية حساب الأوزان كما جاء في بحث Zahedi 1993. وأن أوزان التعلم تعتبر عامل هام في تحديد قيم مدخلات النيورون وتؤثر أيضا بطريقة غير مباشرة علي المخرجات، ونجد أن هناك نوعين من التعلم هما:

- التعلم الذاتى.
- التعلم الغير موجه.

إن طريقة التعلم الذاتى تتطلب توافر قيم المدخلات وقيم المخرجات المطلوبة ويقوم نظام الشبكات العصبية باستقبال المدخلات الحقيقية، وحساب الخطأ ثم تعديل الأوزان لقيمة الخطأ المحسوب.

اما طريقة التعلم الغير موجه فإن المخرجات غير معلومة ولكن المدخلات معلومة وبالتالي لا يتم تعديل الأوزان وهذا النوع من التعلم يعتبر مناسب في اسلوب التحليل العنقودي.

ونجد أن الشبكات العصبية يمكن أن تمر بثلاث مراحل هي:

1 – مرحلة التعلم (التدريب):

إن الشبكة العصبية يتم تدريبها باستخدام فئة بيانات التدريب، ثم يتم تعديل الأوزان وذلك لتدنية دالة الهدف (وعلي سبيل المثال فإن دالة الهدف قد تكون جذر مجموع مربعات الخطأ RMS).

2 – مرحلة الاختبار:

يتم اختبار الشبكة باستخدام فئة بيانات الاختبار حيث انه في هذه المرحلة تصبح أوازن الشبكة ثابتة كنتيجة لتطبيق المرحلة الأولي وهى مرحلة التعلم.

3 – مرحلة التشغيل:

يتم تطبيق الشبكة العصبية المقترحة علي مجموعة بيانات نتائجها غير معلومة.

3. 5. 1. قواعد التعلم:

إن قواعد التعلم تمثل المعادلة التي يتم استخدامها لتعديل أوزان الاتصال بين النيورنات. ونجد اكثر قواعد التعلم استخداما هي:

1 – Hebb's rule.

2 – Hopfield rule.

.Delta rule – 3

:Hebb's rule – 1

تعتبر هذه القاعدة أول قاعدة تعلم وأيضا تعتبر من أفضل قواعد التعلم والتي تم تطويرها بواسطة العالم Hebb في كتابة بعنوان تنظيم السلوك والذي تم اصدارة عام 1949 وتتلخص هذه القاعدة فيما يلي:

إذا استقبل نيورون مدخلات من نيورون آخر وكان كليهما (النيورون المستقبل و النيورون المرسل) له نشاط عالي (أي لهما نفس الإشارة الرياضية) فإن أوزان الاتصال يجب أن تزيد والعكس صحيح.

:Hopfield rule – 2

أن هذة القاعده تعتبر قريبة الشبه من قاعدة Hebb's rule ولكن الاختلاف بينهما يتمثل فى أن هذة القاعده تحدد نقاط القوة والضعف فى الاتصالات بين النيورونات، حيث أنه إذا كانت المخرجات الحقيقة والمدخلات كليهما نشط أو كليهما غير نشط فإن أوزان الاتصال تزيد بمقدار يسمي معدل التعلم learning rate. أما في حالة اختلاف طبيعة المخرجات عن المدخلات كأن يكون إحداهما نشط والآخر غير نشط فإن أوزان الاتصال تنقص بمقدار معدل التعلم (Hegde, Sweet, and Levy, 1987).

:Delta rule – 3

إن طريقة الارتداد الخلفي المعياري Standard backprop تعرف بأسم قاعدة دالتا المعممة *generalized delta rule*، ولقد تم تطوير هذه الطريقة للتعلم بواسطة Rumelhart, Hinton, and Williams في عام 1986. وتعتبر هذه الطريقة من اكثر الطرق استخداما في حالة التعلم الذاتى. أن قاعدة دلتا للتعلم التي تتميز بوجود ما يسمي بعامل الدفع Momentum

Factor (انظر الفقرة 3.5.2. (2)) تسمي نظرية الكرة الثقيلة في طرق التحليل العددي (Polyak, 1987; Bertsekas 1995). إن هذه القاعدة تعتمد علي التعديل المستمر لأوزان الاتصال لتقليل الفرق (دلتا) بين المخرجات الحقيقية ومخرجات الشبكة. هذه الطريقة تقوم بتغيير أوزان الاتصال حتى تصل الشبكة إلي أقل متوسط مجموع مربعات الخطأ للشبكة. ويمكن الإشارة إلي هذه الطريقة للتدريب أو التعلم باسم قاعدة متوسط مجموع المربعات للتعلم.

إن الطريقة التي تعمل بها هذه القاعدة هي أن فرق الخطأ في طبقة المخرجات يتم تحويله بواسطة تفاضل دالة التحويل، ثم يتم استخدامه في الطبقة السابقة لتعديل أوزان اتصال المدخلات وبطريقة أخري فأن الخطأ يتم إعادته (للخلف) إلي الطبقات السابقة طبقة في كل مرة ويتم تكرار عملية الإرجاع حتى نصل إلي الطبقة الأولي.

3. 5. 2. معلمات التعلم:

تعتبر معلمات التعلم الأداة التي تستخدم لتحسين أداء الشبكة، وهناك ثلاث معلمات للتعلم هي:

1 – معدل التعلم Learning Rate.

2 – عامل الدفع Momentum factor .

3 – حد التجاوز عن خطأ التدريب Training Tolerance.

1 – معدل التعلم:

معدل التعلم هو حدود تعديل أوزان الاتصال فإذا كان معدل التعلم كبير فهذا يؤدي إلي عدم ثبات الشبكة وإذا زاد معدل التعلم بقدر كبير فإن قدرة الشبكة علي التنبؤ تصبح منخفضة. أما إذا كان معدل التعلم صغير للغاية فهذا يعتبر

ضياع للوقت حيث تستغرق الشبكه وقت اكثر لإجراء عملية التعلم (Garson, 1991). ونجد أن زيادة معدل التعلم يعتبر إجراء مفيد لزيادة قدرة الشبكة علي التعلم وبالتالي يصبح من الممكن تعديل الأوزان.

2 – عامل الدفع Momentum Factor:

هذا العامل يمثل نسبة التحيز في الأوزان من مرحله إلي أخري. ونجد إن انخفاض قيمة هذا العامل تؤدي إلي عدم ثبات الأوزان وتعوق الشبكة من اجراء عملية التعلم. أما في حالة زيادة قيمة هذا العامل فإن تعديل الشبكة يصبح عمليه صعبه، وللحصول علي شبكة عصبية ثابتة يجب أن يكون هذا العامل أقل من واحد صحيح. ولقد ذكر Garson في بحثه عام 1998 أن اقتراب قيمة هذا العامل من الواحد الصحيح يعتبر مطلوب حتى يمكن تمهيد تقلبات الخطأ.

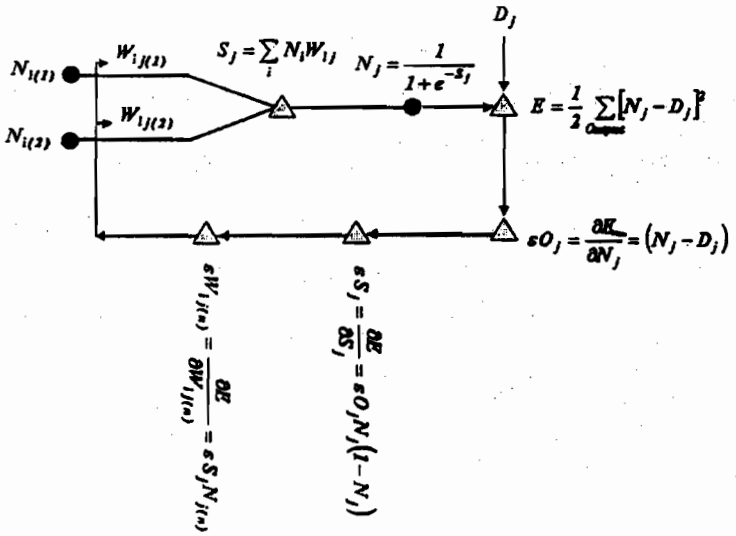
3 – حد التجاوز عن خطأ التدريب Training Tolerance:

إن هذه القيمة تمثل الخطأ الهامشي أو الخطأ المسموح به أثناء المقارنة بين مخرجات الشبكة والمخرجات الحقيقة فإذا كان قيمة هذا العامل تساوي صفر فإن هذا يعني أن مخرجات الشبكة يجب أن تتطابق مع المخرجات الحقيقية. أما إذا زادت قيمة هذا العامل فإن هذا يعني انخفاض دقة التنبؤات بواسطة الشبكة. ويتم تحديد قيمة هذا العامل بالمحاولة والخطأ واعتماداً علي خبرة الباحث وطبيعة البيانات.

4 – أسلوب الارتداد الخلفي Backpropagation algorithm:

إن أسلوب الارتداد الخلفي أو التغذية الخلفية أحد الطرق التي أدت إلي استخدام أسلوب الشبكات العصبية في تطبيقيات كثيره مختلفة. إن هذا الأسلوب تم تطويره بواسطة Rumelhart, Hinton, Williams فى عام 1986، حيث قاموا ببناء أول شبكة عصبية لها أكثر من طبقة خفيه واحدة. ونجد إن الهدف من استخدام هذا الأسلوب هو حساب الخطأ في طبقة المخرجات ثم إعادته مرة أخري لكل طبقة خفية لتعديل أوزان الاتصال حتى يتم الوصول إلي طبقة المدخلات (Dhar and Stein, 1996). الرسم التالي (شكل رقم 1) يوضح شبكة عصبية صناعية ذات نيورون واحد والتي يتم تدريبها باستخدام أسلوب الارتداد الخلفي.

شكل رقم (1) شبكة عصبية ذات نيورون واحد يتم تدريبها باستخدام أسلوب الارتداد الخلفي.



إن أسلوب الارتداد الخلفي يهدف إلي تدنية عنصر الخطأ والذي يتمثل في الفرق بين مخرجات الشبكة والمخرجات الحقيقية. بعد حساب الخطأ بين مخرجات الشبكة والمخرجات الحقيقية، يتم تغذية هذا الخطأ في اتجاه عكسي (عكس اتجاه المدخلات) للشبكة لتعديل الأوزان وبالتالي يتم تخفيض الخطأ (Fahlmann, 1988) هذه الإجراءات يتم اعادتها حتى يصل الخطأ إلي أقل قيمة ممكنة.ويمكن استخدام المعادلة رقم (1) لتعديل الأوزان W_{ij} من العقدة (النيورون) N_i إلي العقدة أو النيورون N_j:

$$W_{ij,(t+1)} = W_{ij,t} + (\lambda)(\varepsilon W_{ij})(N_i) \qquad (1)$$

حيث أن:

W_{ij} = تمثل وزن الاتصال بين النيورون N_i والنيورون N_j

t = تمثل عدد مرات تعديل الأوزان في الشبكة.

λ = تمثل معدل التعلم (أحد معلمات التعلم في الشبكة)

εW_{ij} = تمثل حساسية النيورون N_j لتغير الوزن W_{ij}

يلي ذلك حساب إجمالي المدخلات للنيرون N_j كما يلي:

$$S_j = \sum_i N_i W_{ij} \qquad (2)$$

حيث أن:

S_j = تمثل مجموع جميع المدخلات لنيرون

N_i = تمثل مخرجات النيورون السابق

W_{ij} = تمثل وزن الاتصال بين النيورون i والنيورون j في الطبقة السابقة.

يلي خطوة حساب إجمالي المدخلات للنيورون i تحويل المخرجات باستخدام دالة التنشيط أو التحويل اللوجستي والتي يمكن صياغتها كما يلي:

$$N_j = \frac{1}{1+e^{-S_j}} \qquad (3)$$

حيث أن:

N_j = تمثل إجمالي المخرجات من النيورون j

ونجد أن الخطأ الكلي للبيانات التي مرت خلال الشبكة يمكن حسابة كما يلي:

$$E = \frac{1}{2}\sum_{Output}[N_j - D_j]^2 \qquad (4)$$

حيث أن:

D_j = تمثل المخرجات الحقيقية للنيرون j

هذا الخطأ الذي تم حسابه يتم إرجاعه إلي الشبكة في اتجاه عكسي حتى يتم تخفيض الخطأ وهذه العملية يمكن تنفيذها من خلال 4 خطوات هي:

الخطوة الأولي:

يتم حساب الخطأ لمخرجات النيورون والتي يرمز لها بالرمز O_j ونجد أن هذا الخطأ يمكن حسابه باستخدام المعادلة التالية:

$$\varepsilon O_j = (N_j - D_j) \qquad (5)$$

الخطوة الثانية:

إن مقدار الخطأ يتغير كلما تغيرت مدخلات النيورون لمخرجات نيورون آخر وهذا المقدار يجب حسابه ويمكن تحديد هذا المقدار بمقدار التغير في المعادلة رقم (4) عندما يتغير إجمالي المدخلات في المعادلة رقم (2) والتي يمكن التعبير عنها رياضيات كما يلي:

$$\varepsilon S_j = \varepsilon O_j N_j (1 - N_j) \qquad (6)$$

الخطوة الثالثة:

يتم حساب مقدار تعديل الأوزان W_{ij} من النيورون N_i فى الطبقه التى تقع اسفل الطبقه الحاليه الى النيورون N_j الموجود فى الطبقه الحاليه كما يلي:

$$\varepsilon W_{ij} = \varepsilon S_j N_j \qquad (7)$$

الخطوة الرابعة:

هذه العملية يتم إعادتها من النيورون في الطبقات الدنيا وذلك بالسماح للنيورنات في الطبقة الخفية لعرض مخرجات الطبقة. ثم يتم حساب إجمالي الخطأ من جميع المدخلات إلي الطبقة الخفية. هذا بالإضافة إلي أنه يتم حساب مقدار الخطأ في النيورون الخفي عن طريق حساب التغير في النيورون في الطبقة التي تقع أعلي الطبقة الخفية. ويمكن حساب مقدار تعديل الأوزان باستخدام المعادلة التالية:

$$\varepsilon H_i = \sum_j \varepsilon S_j W_{ij} \qquad (8)$$

حيث أن:

H_i= هو الفرق بين وزن الاتصال في خطوة ووزن الاتصال في الخطوة التالية والتى يمكن حسابه بأستخدام المعادله التاليه:

$$H_i = W_{ij(1)} - W_{ij(2)}$$

ولقد أشار Dhar and Stein فى بحث لهما تم نشره عام 1996 إلي أن الخطأ في الشبكة يتم إرجاعه عدة مرات حتى يتم تخفيض الخطأ الكلي لمخرجات الشبكة.

ونجد أن مشكلة تجاوز التدريب Over Training يمكن أن تحدث لأي شبكة عصبية، وهذه المشكلة تحدث عندما يتم تدريب الشبكة بطريقة مثلي ولكن تعجز الشبكة عن تعميم النتائج في فئة الاختبار Validation set. ولتفادي هذه المشكلة يمكن استخدام أحد الطرق الآتية أثناء تدريب الشبكة، هذه الطرق هي:

- طريقة Cross Validation.
- طريقة Jackknife.
- طريقة Bootstrap.

4. 1. طريقة Cross Validation:

إن هذه الطريقة تقوم علي استخدام فئة الصلاحية Validation set للبيانات لتحديد متي يتم إيقاف تدريب الشبكة ونجد إن تدريب الشبكة يستمر طالما يتم تحسين الخطأ الناتج من الشبكة. وعندما يتوقف انخفاض الخطأ فإن الشبكة ستتوقف عن التدريب أو التعلم.

إن هذا الإجراء يسمي حفظ الأفضل Save Best وبعد اختيار أفضل شبكة وحفظها، يتم اختبارها مع فئه بيانات جديدة لتحديد إمكانية تعميم النتائج باستخدام الشبكة.

ونجد أنه في طريقة k-Fold Cross Validation يتم تقسيم البيانات إلي k فئة جزئية متساوية الأحجام. وفي كل مرة تقوم الشبكة بالتدريب k مرة، في كل مرة يتم ترك فئة جزئية واحده واستخدام باقي الفئات الجزئيه في التدريب أي يتم (k-1) فئه جزئية ويتم حساب الخطأ من الفئة الجزئية المتروكة.

وإذا كانت k تساوي حجم العينة فإن هذه الطريقة تسمي ترك وحدة واحدة "Leave-one-out". ونجد أن طريقة Leave-ν-out اكثر تعقيداً من

طريقة cross-validation حيث أنه يتم ترك جميع الفئات الجزئية ذات الحجم ν مشاهدة. كما أنه يوجد اختلاف بين طريقة Cross-Validation وطريقة Split-Sample والتي تستخدم فى طريقة التوقف المبكر للشبكات العصبية. ففي طريقة Split-Sample يتم استخدام فئة جزئية واحدة لتقدير الخطأ المعمم بدلا من k فئة جزئية مختلفة. ونجد أن (Goutte 1997) قد توصل إلي أن طريقة Cross Validation تعتبر طريقة صالحة للبيانات الصغيرة وذلك في الرد علي البحث الذي قام كل من Zhu and Rohwer في عام 1996 والذي توصلا فيه إلي أن طريقة Cross-Validation تصلح للبيانات التي حجمها كبير فقط.

4. 2. طريقة Jackkknife:

لقد لوحظ وجود تشابه بين طريقة Leave-one-out cross-validation وطريقة jackknife، حيث أن كليهما يعتمد علي حذف كل مشاهدة بالترتيب والتتابع وإعادة تدريب الشبكة علي المشاهدات الباقية. ولكن الطريقة الأولي تستخدم لتقدير الخطأ المعمم، بينما الطريقة الثانية (Jackknife) تستخدم لتقدير تحيز الإحصاء المقدر ففي الطريقة الثانية يتم تقدير الإحصاء المطلوب من كل فئة جزئية ثم حساب متوسط هذه التقديرات. ومقارنة المتوسط بقيمة الإحصاء المقدر باستخدام جميع المشاهدات وبالتالي يمكن أيضا استخدام الطريقة الثانية لتقدير الخطأ المعياري للإحصاء المقدر.

4. 3. طريقة Bootstrap:

إن هذه الطريقة تعتمد على سحب عينات عشوائيه من البيانات، حيث أن كل عينه يتم سحبها من البيانات الأصلية مع الإرجاع، واعادة التحليل لكل عينه عشوائيه بدلاً من تقسيم المشاهدات إلي فئات جزئية وإعادة التحليل لكل فئة

جزئية وقد أشار كل من Efron and Tibshirani في عام 1993 أن عدد العينات يجب أن يتراوح بين 50 عينة و 2000 عينة.

5 – مميزات وحدود الشبكات العصبية:

المميزات:

- إن نماذج الشبكات العصبية تعتبر نماذج غير خطية ويمكن تعديلها وبالتالي يمكن استخدمها في التعرف علي العلاقات الغير خطية بطريقة أكثر كفاءة من نماذج الانحدار المتعدد. (Gorr, Nagin & Szczypula 1994; and Marquez et al., 1991)
- عند تطبيق تحليل الانحدار بأستخدام الشبكات العصبية فأن هذا النوع من التحليل لايتطلب تحديد مسبق لطبيعة العلاقات بين المتغيرات وأيضا لا يتطلب تحديد مسبق لشكل النموذج المقترح. ولقد وجد Sarle في عام 1994 أن الشبكات العصبية تعطي نتائج أفضل في حالة احتواء البيانات علي قيم متطرفة أو في حالة عدم تحديد شكل النموذج المقترح.
- إن استخدام الشبكات العصبية المعتمدة علي الانحدار لا يتطلب توافر فروض استخدام تحليل الانحدار فنجد أن الشبكات العصبية لا تتطلب توافر شرط عدم ارتباط المتغيرات المستقلة وبهذا فأنها تحل مشكلة الارتباط الذاتي وكذلك لا تتطلب أن تكون المخرجات مستقلة عن بعضها البعض وايضاً لا تتطلب توافر شرط اعتدالية البواقى.
- تعتبر الشبكات العصبية أداة تحليل إحصائي قوية وذلك لقدرتها علي التعلم من الخبرة السابقة وتعديل نفسها طبقا للبيانات الجديدة.
- تعتبر الشبكات العصبية أكثر أدوات التحليل الإحصائي قوة في حالة احتواء البيانات علي مفردات شاذة.

حدود الشبكات العصبية:

تعتبر الشبكات العصبية بمثابة الصندوق الأسود فهي قادرة علي التنبؤ بدقة ولكن لا توفر النموذج الذي يمكن استخدامه لشرح طبيعة العلاقات بين المتغيرات وبعضها البعض ولكن يمكن التغلب علي هذه المشكلة باستخدام نماذج الانحدار بالتطبيق علي الشبكات العصبية، كما إن استخدام الشبكات العصبية لا يوفر القدره على استخدام اختبارات الجوهرية المستخدمة في التحليل الإحصائي.

6 – العلاقة بين الشبكات العصبية والطرق الإحصائية:

إن الاستدلال الإحصائي يعني في الشبكات العصبية التدريب أو التعلم من البيانات التي تحتوي علي مفردات شاذة او التى بها تقلبات كثيره، إن معظم الشبكات يمكن تدريبها لاكتشاف المفردات الشاذة وتقلبات البيانات وعلي سبيل المثال:

- إن شبكات التغذية الأمامية بدون طبقة خفيفة تعادل النماذج الخطية المعممه.
- إن شبكات التغذية الأمامية ذات طبقة خفيفة واحدة تعادل الانحدار projection pursuit regression
- إن الشبكات العصبية الاحتمالية تعادل تحليل كيرنال للتمايز kernel discriminant analysis.
- إن شبكات كوهين تعادل التحليل العنقودي ذو عدة أوساط.
- إن طريقة هيبيان للتعلم تعادل طريقة تحليل المكونات الرئيسية.

7 – تطبيقات الشبكات العصبية:

لقد لوحظ زيادة الأبحاث التي تقوم علي المقارنة بين تطبيقات الشبكات العصبية واستخدام أساليب التحليل الإحصائي التقليدية في الفترة الأخيرة. فلقد قدم كل من Hertz, Krogh, and Palmer في عام 1991 دراسة مقارنة بين الشبكات العصبية والإحصاء، وفي عام 1992 قدم Hinton بحث يكشف فيه عن الخلفية الإحصائية في تحليل الشبكات العصبية. أما Weiss and Kulikowski فقد قدما في عام 1991 بحث للمقارنه بين طرق التمايز (التصنيف) لعدد من الشبكات العصبية والنماذج الإحصائية. ولقد لوحظ أن هناك تركيز كبير علي استخدام الشبكات العصبية في المجال الاقتصادي والمالي فقد قدم العديد من الاقتصاديون تطبيقات للشبكات العصبية في المجال الاقتصادي.

ولقد قدم ايضاً العديد من الباحثين أبحاث تستخدم الشبكات العصبية في بيانات الأسواق المالية حيث أن عدم خطية البيانات يمثل صعوبات في التحليل الإحصائي التقليدي. ونجد أن الاستخدام الرئيسى للشبكات العصبيه يقع فى مجال التطبيقات الماليه والاقتصاديه والتى تتميز بياناتها بعدم الخطيه. وهناك العديد من الاقتصاديين قدموا الكثير من الابحاث فى مختلف مجالات الاقتصاد (Kuan and White, 2002; Bierens, 1994). وايضاً هناك العديد من الباحثين قدموا ابحاث لاستخدام الشبكات العصبيه فى مجال سوق المال حيث ان عدم خطية البيانات تمثل عقبه فى اجراء التحليلات الاحصائيه التقليديه (Omerod, 1991).

ولقد قدم Verkooijen فى بحثه المنشور عام 2003 التطبيقات التقليديه والتى يمكن استخدام الشبكات العصبيه فيها، وهذة التطبيقات هى:

- التنبؤ بسعر العملات
- ترتيب السندات
- التنبؤ بفشل الاعمال
- مخاطر الاقتراض
- التنبؤ بسرقات البنوك
- التنبؤ بانهيار البنوك

ونجد أن الشبكات العصبيه يمكن استخدامها كاداه للتحليل بدلاً من الاساليب الاحصائيه التاليه:

التصنيف (التمايز):

حيث يتم تصنيف المفرادات (مشاهدات العينه) الى فئتين اعتماداً على انتماء المفرده الى مجتمع ما، فعلى سبيل المثال يمكن تصنيف قرض البنك على انه قرض جيد أى أن نسبة الخطر فى عدم السداد منخفضه وهناك قرض اخر يتم تصنيفه على أنه قرض سيىء أى أن نسبة الخطر فى سداد القرض مرتفعه (Farag\'o and Lugosi, 2001).

التحليل العنقودى:

يتم تقسيم المشاهدات الى مجموعات كل مجموعه لها نفس الخصائص. مثل تقسيم العملاء طبقاً لنوع المشتريات التى يقوموا بشرائها (Wong and Lane, 2002).

النمذجه:

أن الانسان يمكنه التعلم لمعرفة طبيعة العلاقات بين بعض المتغيرات ولكن اسلوب النمذجه يشرح هذة العلاقات ى صورة نماذج يمكن تعميمها (Rissanen, 2001).

التنبؤ:

فى تحليل السلاسل الزمنيه يمكن التعرف على سلوك السلسله الماضى واستخدام هذا السلوك فى التنبؤ بقيم السلسله فى المستقبل (Husmeier, 1999 and Wong, 1999).

تحقيق القيود والامثليه:

يمكن استخدام الشبكات العصبيه فى تحقيق دالة هدف ما فى ظل تحقق مجموعه من القيود (غالباً ما تكون متعارضه) وهو ما يسمى باسلوب بحوث العمليات (Pearl, 1988; Ripley, 1993).

8. النتائج:

أن الشبكات العصبيه لها العديد من المزايا والتىتفوق طرق التحليل التقليديه، هذة المزايا يمكن سردها كما يلى:

(1) استخدام الشبكات العصبيه لا يتطلب توافر فروض عن طبيعة توزيع البيانات، وبالرغم من توافر هذا الشرط فأنها لا تعتبر طريقة تحليل متحيزه. فنجد أنه بدلا من صياغة فروض عن المجتمع محل الدراسه فأن الشبكات العصبيه ذات طبقه خفيه واحده على الاقل يمكنها تحليل البيانات والتعرف على العلاقات بين المتغيرات (White, 1992).

(2) أن بيانات السلاسل الزمنيه تتميز بأنها بيانات ديناميكيه، ولهذا كان يجب توفير اداه للتعرف على عدم خطية البيانات، وتعتبر الشبكات العصبيه من افضل ادوات التحليل التى يمكنها اكتشاف عدم الخطيه (Wasserman, 1989)

(3) أن الشبكات العصبيه تعتبر افضل ادوات التحليل فى حالة عدم اكتمال البيانات أو فقد بعض البيانات (Kuo and Reitsch, 1995/1996).

(4) يمكن للشبكات العصبيه الحصول على التنبؤات فى وقت اقصر من الوقت الازم لاستخدام النماذج الاقتصاديه القياسيه.

(5) يمكن للشبكات العصبيه تجاوز عجز طرق الانحدار التقليديه فى حالة احتواء البيانات على نسبة تشويش عاليه أو احتواء للبيانات على مفردات شاذه.

من هذة المناقشه يمكن حصر استخدام الشبكات العصبيه فى الاربعة نقاط التاليه:

(1) التنبؤ أو توفيق الدوال

(2) البيانات التى بها نسبة تشويش

(3) التعرف على نمط البيانات واكتشاف العلاقات

(4) التصنيف

ويمكن تحديد احتياج الباحث أو عدم احتياجه الى استخدام الشبكات العصبيه فى النقاط التاليه:

- أذا كان الباحث يرغب فى الحصول على تنبؤات دقيقه ولا يحتاج الى تفسير او شرح لمعلمات النموذج فأن الشبكات العصبيه تعتبر من اقوى ادوات التنبؤ.
- أذا كان الباحث يرغب فى الحصول على تنبؤات دقيقه ويرغب ايضاً فى شرح معلمات النموذج المستخدم فيمكنه تطبيق نماذج الانحدار واسلوب الشبكات العصبيه معاً.

- اذا كانت فروض نماذج الانحدار غير متوافره فيمكن استخدام اسلوب الشبكات العصبيه.
- أذا كانت فروض نماذج الانحدار متوافره فمن الافضل استخدام نماذج الانحدار فى التحليل.

9. ابحاث مقترحه:

- تقييم اسلوب المفاضله بين الشبكات باستخدام طرق بناء الشبكات وطرق التدريب.
- دراسة تأثير طرق التدريب وعدد مرات التدريب على المفاضله بين الشبكات المختلفه.
- دراسة تاثير طبيعة البيانات ومجال التطبيق على المفاضله بين الشبكات المختلفه مع المقارنه بأساليب التحليل الاخرى.
- تطوير قواعد عامه لتصميم وبناء الشبكات.
- تطوير مجموعه من القواعد لتحديد افضل طريقة تعلم أو تدريب للشبكات طبقاً للتطبيقات المختلفه.