Accounting Mining Data Using Neural Networks (Case study)

التنقيب المحاسبي عن البيانات باستخدام الشبكات العصبية (حالة دراسية)

أ.م.د. وحيد محمود رمو/ كلية الادارة والاقتصاد/ جامعة الموصل
Dr.waheedramo@yahoo.com

OPEN ACCESS
P - ISSN 2518 - 5764
E - ISSN 2227 - 703X

Received:20/2/2018
Accepted: 16/10/2018

المستخلص

واجهت منظمات الاعمال في العصر الراهن تحديات عديدة أهمها تقنيات المعلومات وذلك لانتشارها الواسع وسهولة استخدامها، اذ أدى استخدامها الى تضخم حجم البيانات التي تتعامل معها منظمات الاعمال بصورة لم يسبق لها مثيل واصبح حجم البيانات المتاحة من خلال الانترنت مشكلة تسعى اطراف عديدة لإيجاد حلول لها فما هو الجدوى من وجودها بهذا الحجم وبهذه الصورة العشوائية، وكشفت التوقعات أنه في العام 2017 سيكون هناك أجهزة متصلة بالإنترنت تقدر بثلاثة أضعاف سكان الأرض، وفي العام 2015 تم نقل أكثر من مليون ونصف المليون غيغا بايت من البيانات في كل دقيقة عالمياً. من هنا ظهر ما يسمى بالتنقيب عن البيانات Data Mining كتقنية تهدف إلى استنتاج المعرفة من كميات هائلة من البيانات، تعتمد على الخوارزميات الرياضية والتي تعد أساساً للتنقيب عن البيانات وهي مستمدة من العديد من العلوم مثل علم الإحصاء والرياضيات والمنطق وعلم التعلم، والذكاء الاصطناعي والنظم الخبيرة، وعلم التعرف على الأنماط وغيرها من العلوم والتي تعتبر من العلوم الذكية وغير التقليدية. وتجسدت مشكلة البحث في ان الازدياد المضطرد في حجم البيانات فضلاً عن ظهور مجالات انية عديدة تحتاج الى بيانات مختلفة نتيجة البيئة المعاصرة لمنظمات الاعمال اليوم يجعل نظم المعلومات غير قادرة على تلبية احتياجات هذه المنظمات الانية وهذا ينطبق تماما على نظم المعلومات المحاسبية كونها النظام الرئيس في منظمات الاعمال اليوم، فان هذه النظم قد تم تصميمها لتلبية حاجات محددة جعلها تعجز اليوم عن تلبية الاحتياجات المختلفة وفق البيئة المعاصرة لمنظمات الاعمال اليوم، فضلاً عن عجزها في التعامل مع حجم البيانات المضطرد الذي اوجدته تقنيات المعلومات. واطلق البحث من فرضيتين اساسيتين الاولى ان اعتماد التنقيب المحاسبي عن البيانات يؤدي الى توفير بيانات لم يكن نظام المعلومات المحاسبي قادر على توفيرها فضلاً عن اختصار الوقت والجهد اللازم في الحصول عليها. بينما كانت الثانية في ان اعتماد التنقيب المحاسبي عن البيانات يمكن من اعتماد أساليب الذكاء الصناعي في معالجة تلك البيانات لتوفير معلومات مفيدة لترشيد القرارات. وقد توصل البحث الى مجموعة من الاستنتاجات منها ان التزايد المضطرد بحجم البيانات عموماً ولاسيما البيانات المحاسبية يجعل من التعامل وفق الاطر التقليدية مسالة صعبة جداً كما انه يؤدي الى ضياع الوقت والجهد في استخلاص المعلومات. وان ظهور كثير من المتغيرات الانية نتيجة المتغيرات في بيئة العمل يتطلب وجود ادوات تقنية لديها المرونة الكافية للتعامل معها. وكذلك ان ادوات التنقيب عن البيانات لديها القدرة في استنتاج علاقات بناء على قواعد البيانات الموجودة لديها لم تكن موجودة في السابق. توصل البحث الى مجموعة من التوصيات اهمها ضرورة اعتماد النموذج الذي قدمه البحث وهو شبكة (Multilayer Perceptron) وتوجد هذه الشبكة ضمن برنامج (SPSS) الذي يتيح امكانية استخدام هذه الشبكة بسهولة في ترشيد قرار اختيار المشاريع المنفذة في مجالس المحافظات

المصطلحات الرئيسة للبحث/ التنقيب المحاسبي، الشبكات العصبية الاصطناعية، شبكة (Perceptron Multilayer) ، تدريب الشبكة العصبية

Journal of Economics and Administrative Sciences
2019; Vol. 25, No.111
Pages:531 549

المقدمة

واجهت منظمات الاعمال في العصر الراهن تحديات عديدة أهمها تقنيات المعلومات وذلك لانتشارها الواسع وسهولة استخدامها، اذ أدى استخدامها الى تضخم حجم البيانات التي تتعامل معها منظمات الاعمال بصورة لم يسبق لها مثيل واصبح حجم البيانات المتاحة من خلال الانترنت مشكلة تسعى اطراف عديدة لإيجاد حلول لها فما هو الجدوى من وجودها بهذا الحجم وبهذه الصورة العشوائية، وكشفت التوقعات أنه في العام 2017 سيكون هناك أجهزة متصلة بالإنترنت تقدر بثلاثة أضعاف سكان الأرض، وفي العام 2015 تم نقل أكثر من مليون ونصف المليون غيغا بايت من البيانات في كل دقيقة عالمياً.

وأظهرت بيانات معلوماتية نشرتها شركة إنتل الأميركية أن العالم أنشأ بيانات ورفعها على الإنترنت في العام 2013 تقدر بنحو 4 زيتا بايت أو ما يعادل 4 ضرب 10 مرفوعا للقوة 21. يشار إلى أن كل زيتابايت يعادل مليار تيرا بايت ويساوي 10 مرفوعة للقوة 21 بايت.

وبحسب البيانات والتوقعات، فقد شكل موقعا نتفليكس ويوتيوب أكثر من نصف الحركة على الإنترنت، حيث، تتم مشاهدة 138889 ساعة فيديو على موقع يوتيوب كل ساعة، بينما يتم رفع 100 ساعة فيديو على الموقع نفسه، في حين أنه تتم مشاهدة 23148 ساعة فيديو على موقع نتفليكس، ويتم الاستماع إلى 31733 ساعة موسيقى في موقع باندورا.

وأن 4.1 مليون عملية بحث تتم عبر موقع غوغل كل دقيقة، في حين يتم مشاركة 3.3 مليون "قطعة" عبر فيسبوك، سواء أكانت صورة أو مادة مكتوبة أو غيرها، بينما يتم إرسال 347222 تغريدة عبر موقع تويتر، ويتم تحميل 38194 صورة على موقع إنستاغرام.

وهنا نتساءل ما أهمية هذه البيانات في ظل وجود حقيقة تشير إلى أن المعلومات المنظمة من هذه البيانات لا تشكل إلا جزءً ضئيلاً يصل إلى 10% مقارنه بالمعلومات غير المنظمة والتي تشكل الباقي، ان كل ذلك ادي إلى ازدياد الحاجة إلى تطوير أدوات تمتاز بالقوة لتحليل البيانات واستخراج المعلومات والمعارف منها، فالأساليب التقليدية والإحصائية لا تستطيع أن تتعامل مع هذا الكم الهائل لذا تستخدم أدوات ذكية لمعالجة هذه البيانات.

من هنا ظهر ما يسمى بالتنقيب عن البيانات Data Mining كتقنية تهدف إلى استنتاج المعرفة من كميات هائلة من البيانات، تعتمد على الخوارزميات الرياضية والتي تعد أساساً للتنقيب عن البيانات وهي مستمدة من العديد من العلوم مثل علم الإحصاء والرياضيات والمنطق وعلم التعلم، والذكاء الاصطناعي والنظم الخبيرة، وعلم التعرف على الأنماط وغيرها من العلوم والتي تعد من العلوم الذكية وغير التقليدية.

مشكلة البحث:

ان الازدياد المضطرد في حجم البيانات فضلاً عن ظهور مجالات انية عديدة تحتاج الى بيانات مختلفة نتيجة البيئة المعاصرة لمنظمات الاعمال اليوم يجعل نظم المعلومات غير قادرة على تلبية احتياجات هذه المنظمات الانية وهذا ينطبق تماما على نظم المعلومات المحاسبية بوصفها النظام الرئيس في منظمات الاعمال اليوم، اذ ان هذه النظم قد تم تصميمها لتلبية حاجات محددة جعلها تعجز اليوم عن تلبية الاحتياجات المختلفة على وفق البيئة المعاصرة لمنظمات الاعمال اليوم، فضلاً عن عجزها في التعامل مع حجم البيانات المضطرد الذي اوجدته تقنيات المعلومات.

ويمكن تحديد مجالات عجز هذه النظم كما يأتي:

1. صعوبة تحديد أنواع البيانات المطلوب جمعها ومصادر الحصول عليها، في ضوء حجم البيانات الهائل المتاح وكذلك في ضوء وجود مجالات انية لم تكن موجودة سابقا.
2. صعوبة التعرف على مختلف المتغيرات المحيطة بالبيانات المطلوبة وتلك المتوقعة، وعلى كيفية تأثيرها في عملها

وتأسيسا على ما تقدم يمكن صياغة التساؤلات البحثية الاتية:

1. هل أن وجود كم هائل من البيانات والمعلومات المحاسبية يتطلب استخدام تقنيات التنقيب عن البيانات لتوفير المعلومات المحاسبية المفيدة لترشيد القرار؟
2. هل أن استخدام تقنيات التنقيب المحاسبي عن البيانات سيؤدي الى تعزيز جودة المعلومات المحاسبية؟

أهمية البحث :

تتضح أهمية البحث جلياً في الازدياد الهائل للبيانات المتاحة، فضلاً عن وجود احتياجات مختلفة اوجدتها البيئة المعاصرة لمنظمات الاعمال، وكذلك ان التعامل مع هذا الحجم من البيانات وفق الانماط التقليدية يؤدي الى استنزاف الوقت وازدياد الكلف.

اهداف البحث:

يهدف البحث الى تسليط الضوء على:

1.مشكلة التزايد المضطرد لحجم البيانات بشكل يجعل من الصعوبة بمكان التعامل معها لانها تستنزف الكثر من الوقت والجهد وتؤدي الى كلف عالية في الحصول على المعلومة المفيدة.

2.ان تصميم نظم المعلومات المحاسبية الحديثة يجب ان يراعي مشكلة الازدياد المضطرد للبيانات التي اوجدتها تقنية المعلومات فضلاً عن الحاجات الانية التي اوجدتها بيئة الاعمال المعاصرة.

فرضية البحث:

يستند البحث الى الفرضيتين الاساسيتين الاتيتين:

1.ان اعتماد التنقيب المحاسبي عن البيانات يؤدي الى توفير بيانات لم يكن نظام المعلومات المحاسبي التقليدي قادر على توفيرها بشكل علمي موضوعي مقبول من كل الاطراف فضلاً عن اختصار الوقت والجهد اللازم في الحصول عليها.

2.ان اعتماد التنقيب المحاسبي عن البيانات يمكن من اعتماد أساليب الذكاء الصناعي في معالجة تلك البيانات لتوفير معلومات مفيدة لترشيد القرارات.

خطة البحث:

لغرض تحقيق اهداف البحث تم تقسيمه الى ثلاث مباحث تناول الأول الإطار المفاهيمي للتنقيب عن البيانات المحاسبية بينما عرض الثاني الاطار المفاهيمي للشبكات العصبية وعرض الثالث استخدام التنقيب المحاسبي عن البيانات في تصميم برنامج لترشيد قرار المفاضلة بين المشاريع الاستثمارية باستخدام الشبكات العصبية (حالة دراسية)

المبحث الأول/ الإطار المفاهيمي للتنقيب عن البيانات المحاسبية

اولاً: مفهوم التنقيب عن البيانات

عرف (Lin, 2007,2) التنقيب عن البيانات بانها عمليه تحليل البيانات من وجات نظر مختلفه وتلخيصها في معلومات مفيده هذه المعلومات يمكن استخدامها لزيادة الإيرادات، وخفض التكاليف، أو كليهما. وان برمجيات التنقيب عن البيانات هي واحده من عدد من الادوات التحليلية لتحليل البيانات. وهو يسمح للمستخدمين بتحليل البيانات من العديد من الابعاد أو الزوايا المختلفة، وتصنيفها ، وتلخيص العلاقات التي تم تحديدها من الناحية الفنية، وتتم عملية استخراج البيانات بالعثور علي علاقات أو أنماط بين عشرات من الحقول في قواعد البيانات العلائقية الكبيرة.

في حين عرفها (العلي واخرون ، 2011 ، 7) انها اجراءات تحليلية دقيقة وذكية تفاعلية وتسلسلية تسمح لمسيري النشاط عند استخدام هذه الاجراءات باتخاذ قرارات والقيام باعمال في صالح النشاط المسؤولين عنه والمؤسسة التي يعملون بها.

ويرى الباحث ان التنقيب في قواعد البيانات يهدف إلى استخلاص المعلومات المخبأة فيها، فاستخدامها يوفر للمؤسسات في جميع المجالات استكشاف والتَّركيز على أهم المعلومات اذ ان التعامل مع حجم كبير من البيانات تظهر مسائل جديدة مثل كيفية تحديد النقاط المميزة في البيانات، وكيفية تحليل البيانات في فترة زمنية معقولة وكيفية قرار ما إذا كانت أي علاقة ظاهرية تعكس حقيقة في طبيعة البيانات وعادة يتم التنقيب في بيانات تكون جزءاً من كامل البيانات حيث يكون الغرض عادة تعميم النتائج على كامل البيانات.

أصبح موضوع التنقيب عن البيانات، من المفاهيم الحديثة في نجاح نظم دعم القرار في منظمات الأعمال ولغايات توفير معلومات سليمة لعملية اتخاذ القرارات اذ برزت الحاجة إلى مرحلة إجرائية لاستخلاص مواصفات وعلاقات من المعطيات وتقديم معلومات جديدة لم تكن معروفة مسبقاً في نظم المعلومات التقليدية سميت بالتنقيب عن البيانات، فأصبح دور نظم المعلومات في هذه البيئة الجديدة ممارسة الدور المعرفي كنظم متكاملة يضيف الى ادوارها التقيدية ادواراً جديدة تعمل على توفير معلومات ومعارف واسعة تساعد على ترشيد القرارات بشكل افضل ويتاتى هذا الدور من :

(Efraim & Dorothy, 2008,183-186)

1.ايجاد وتامين المعرفة (Creating Knowledge) تسعى نظم المعلومات المعرفية لتجهيز العاملين في الحقل المعرفي بالرسومات، والتحليلات، والاتصالات، ووسائل إدارة الوثائق، فضلا عن الوصول إلى مصادر المعلومات والمعرفة الداخلية والخارجية.

2.اكتشاف وتصنيف المعرفة (discovering and codifying knowledge) اذ وفرت نظم الذكاء الصناعي امكانية استنباط ودمج الخبرات لغرض ايجاد نماذج وعلاقات في كميات كبيرة من البيانات وتقوم نظم دعم القرار بتحليل قواعد بيانات واسعة وتستطيع ايضاً اكتشاف معارف جديدة.

3.المشاركة بالمعرفة (Knowledge Sharing) نظم التعاون الجماعية تستطيع ان تساعد العاملين في الوصول والعمل في آن واحد على الوثيقة نفسها ومن مواقع مختلفة ومن ثم التنسيق بين نشاطاتهم المختلفة

4.توزيع المعرفة (Knowledge Distributing) نظم ادوات المكتب وادوات الاتصال تستطيع تامين الوثائق والاشكال الاخرى من المعلومات وتوزيعها على العاملين في مجال المعلومات والمعرفة من خلال ربط المكاتب بوحدات الاعمال داخل الممنظمة وخارجها.

ومن ثم فان المزايا التي توفرها تقنيات التنقيب في البيانات في توفير معلومات دقيقة وصحيحة وبشكل سريع تبين اهمية هذه التقنية في اتخاذ القرار ومن اهم هذه المزايا (حسين، 2004 ، 24)

1.تساعد متخذ القرار على تفعيل الترابط بين الاقسام المختلفة والاعمال المختلفة وينعكس ذلك على مصلحة العمل في المنظمة

2.تسهل التعامل مع تقنيات المعلومات المتطورة وتساعد على قياس فاعلية وانتاجية نظم المعلومات الفرعية المختلفة من خلال توفير المعلومات الدقيقة والصحيحة

3.تساعد في الاستخدام الفعال لمصادر البيانات والموارد المتاحة وفي التخطيط لتحسين وتطوير نظم المعلومات المحاسبية والمصرفية المستخدمة

4.تمكن متخذ القرار في ان يخطط لاستمرارية اجراء الفحص وتحديد المشكلة وعناصرها واتمام الرقابة على ملفات المعلومات والبيانات اللازمة لقراراته ويسعى الى تطوير نظم المعلومات المحاسبية بتكاليف مناسبة

5.تساعد متخذ القرار فيما يحتاج اليه من معلومات لاتخاذ قرارات تسهم في تسريع العمل وانجاز المهام وتبسيط الاجراءات

وهكذا فان بيئة تقنية المعلومات قد اوجدت مفهوماً حديثا لم يكن معمول به سابقا نظرا لما اليه نشاط المالي والاقتصادي من حاجة كبيرة الى المعلومات والى توفير جميع اشكال الربط والاتصال فيما بينها للحصول على معلومات ذات خصائص ودلائل لم تكن بالاهمية التي اصبحت عليه في الوقت الحاضر نظراً لتطور مفاهيم المعرفة

ثانياً: انواع التنقيب عن البيانات وادواته

هناك انواع مختلفة من التنقيب عن البيانات وان إختيار اي منها يعتمد على طبيعة البيانات تحت الدراسة وعلى حجم هذه البيانات وحدد (الدوري، واحمد، 2007 ، 44) وجهات النظر في تبويب انواع التنقيب عن البيانات الى:

1.<u>بحسب طبيعة التنقيب : يقسم الى</u>

أ. الاكتشاف: وهو عملية النظر في قاعدة البيانات لايجاد النماذج من دون ان تكون فكرة محددة عن ما قد تكون عليه.

ب. النمذجة التنبؤية: تستخدم فيه النماذج المكتشفة من قاعدة البيانات للتنبؤ بالمستقبل اي تخمين القيم غير المعروفة بالاعتماد على نماذج سابقة مكتشفة من قاعدة بيانات

ت. التحليل المبرهن: وهو عملية تطبيق النماذج المستخلصة لايجاد عناصر البيانات الشاذة او غير العادية اي انه يبحث في حالات محددة وغير عادية

2.بحسب طبيعة التعامل مع البيانات :

أ. التنقيب العارض: اي النظر الى البيانات من عارض معين وينجز هذا النوع من قبل المحللين.

ب. التنقيب الاستراتيجي: وفيه يكون النظر الى مجموعات اكبر من البيانات الكلية للحصول على فهم شامل لمقاييس مثل الربحية.

ت. التنقيب المستمر: ويحاول فهم كيفية تغير العالم ضمن مدة محددة من الزمن ومحاولة فهم العوامل التي سببت التغير.

ان الادوات التي تستخدمها تقنيات التنقيب عن البيانات عديدة ومتنوعة ولكل واحدة منها دور يخدم غرض معين ومن هذه الادوات: (Crows,1999:10-15):

1. اشجار القرار (Decision Tree)

وهي مشتقة من الاحصاء والذكاء الصناعي وتجد الارتباطات في البيانات وتستخدم في الاستدلال على قواعد العمل وتعد اساس بناء النموذج التنبؤي كما يمكن ان تستخدم اشجار القرار في تحديد المتغيرات الواجب استخدامها كمدخلات للشبكات العصبية.

2.الشبكات العصبية (Neural Network) :

تتكون الشبكة العصبية من طبقات تبدأ بطبقة المدخلات حيث متغير التنبؤ وترتبط عقد المدخلات بعدد من عقد الطبقة الخفية التي ترتبط بدورها بعقد طبقة خفية اخرى او بطبقة المخرجات حيث تضم واحد او اكثر من المتغيرات التابعة.

3.السلاسل الزمنية (Time Series) :

4.وتتنأ بالقيم المستقبلية غير المعروفة بالاعتماد على سلاسل تغير الزمن للمتنبأت فيؤخذ بالحسبان الخواص المميزة كتدرج المدد والموسمية مثلاً

5.استنتاج القاعدة (Rule Induction) :

فيها يتم اشتقاق مجموعة من القواعد المستقلة وعلى خلاف اشجار القرار فهي لا تاتي من شجرة وقد لا تغطي القواعد المكونة كل الحالات الممكنة كما انها تتعارض في تبؤاتها

6.المجاور الاقرب (K-Nearest Neighbor) :

يستخدم المجاور الاقرب فكرة ان حل المشاكل الجديدة يكون عن طريق ملاحظة حلول مشاكل مشابهة تم حلها سابقا.

ويرى الباحث ان اعتماد اي من انواع التنقيب عن البيانات او اي من ادواته يعتمد بالاساس على طبيعة المشكلة فضلاً عن حجم بياناتها والهدف المراد الوصول اليه.

ثالثاً : مجالات استخدام التنقيب عن البيانات في المحاسبة

تتجمع احجام كبيرة من البيانات المحاسبية لدى الكثير من الجهات مثل البورصة ودوائر الضريبة والمصارف ومكاتب تنظيم وتدقيق الحسابات الكبيرة وغالباً تكون كل هذه البيانات متاحة للجمهور، ان استخدام ادوات التنقيب عن البيانات المحاسبية يمكن ان يساهم في حل مشاكل التصنيف والتنبؤ ويسهل من عمليه صنع القرار، ومن المجالات التي يمكن ان يستخدم بها التنقيب عن البيانات المحاسبية هي التنبوء بإفلاس الشركات، وتقدير مخاطر الائتمان، والإبلاغ المستمر، والضائقة المالية، والتنبؤ بأداء الشركات. وقد اعترفت منظمات عديده باهميه التنقيب عن البيانات في مجالي المالية والمحاسبة اذ حدد المعهد الأمريكي للمحاسبين القانونيين التنقيب عن البيانات بوصفه اول عشرة تقنيات مستقبلية، وأدرج معهد مراجعي الحسابات الداخليين التنقيب عن البيانات كاحد الأولويات البحثية الأربع (Koh,2004,462)

ويرى الباحث انه وبسبب قدرة التنقيب عن البيانات في مجال التصنيف والتنبؤ واكتشاف العلاقات بين البيانات وغيرها من المزايا فانه يمكن استخدام التنقيب عن البيانات في المحاسبة وتدقيق الحسابات في المجالات التالية:

1.الاجراءات التحليلية: عملية تدقيق الحسابات فمثلاً تضمن معيار التدقيق 520 عن الاجراءات التحليلية و معيار تدقيق الحسابات (SAS 56) الصادر عن المعهد الامريكي للمحاسبين القانونين تحديد للمقصود بالإجراءات التحليلية بانها النسب والمؤشرات المهمة ومن ضمنها نتائج البحث للتقلبات والعلاقات التي تكون متعارضة مع المعلومات الاخرى ذات العلاقة او تلك التي تنحرف عن المبالغ المتنبأ بها. واكدا ان الإجراءات التحليلية تشكل جزءاً هاماً من عمليه مراجعه الحسابات وتتراوح الإجراءات التحليلية بين المقارنات البسيطة واستخدام النماذج المعقدة التي تنطوي علي العديد من العلاقات وعناصر البيانات. والفرضية الاساسيه التي يقوم عليها تطبيق الإجراءات التحليلية هي ان العلاقات المعقولة بين البيانات قد تفضي إلى العكس. ويتضح جلياً التوافق بين جوهر عملية التنقيب عن البيانات ومجال الاجراءات التحليلية في مهنة تدقيق الحسابات

2.تحديد مخاطر التدقيق: ان تبني مفهوم المخاطر التدقيق يعد من الاتجاهات الحديثة في تدقيق الحسابات، الذي يؤكد على ضرورة تحقيق الأهداف الاستراتيجية للمنظمة وفي هذا النهج يتفهم مدقق الحسابات الأهداف الاستراتيجية ويمكن لادوات التنقيب عن البيانات ان تيسر هذا النهج الجديد لتدقيق الحسابات القائم علي المخاطر (Calderon ,Cheh, 2002,203)

3.التنبوء بفشل الشركات: ان معظم معايير التدقيق ترى ان من مهام مدقق الاحسابات ان يفصح عن التنبوء بفشل الشركات ان توفرت لديه مؤشرات لذلك وبالتاكيد يمكن الافادة من ادوات التنقيب عن البيانات في التوصل الى هكذا مؤشرات.

4.الكشف عن الغش والاحتيال: يشكل الكشف عن الغش والاحتيال مجال مهم يمكن ان تستخدم في ادوات التنقيب عن البيانات اذ يشمل الغش والاحتيال سرقه الأصول فضلا عن اي محاولة (محاولات) لإخفاء السرقة وفي حين ان السرقة المادية للأصول يتم الكشف عنها عاده من خلال أنشطه الفحص والموازنة الروتينية، فان فعل الإخفاء هو الذي يميز الاحتيال عن السرقة العادية وبسبب إخفاء السرقة من المرجح ان يرتكب مرتكب الجريمة نفس الجرائم مراراً وتكراراً، وهناك أنواع مختلفه من الاحتيال بدءاً من الاحتيال بالرعاية الصحية ، والاحتيال بالرهن العقاري، والتلاعب بالتقديرات المحاسبية بالاضافه إلى تضخيم الإيرادات، وتخفيض النفقات الحكومية فضلاً عن الاحتيال في الميزانية الذي يهدف إلى إظهار ديون والتزامات اقل من الالتزامات الفعلية للشركة.وهكذا يمكن استخدام التنقيب عن البيانات في الكشف عن انواع الاحتيال لانه يهدف الى البحث عن الاتجاهات والأنماط في البيانات المحاسبية التي تكشف عن الغش، ويمكن للمنظمات ان تعتمد الخطوات الاتية عند استخدام احد ادوات التنقيب عن البيانات للكشف عن الغش والاحتيال:

1.جمع وفهم البيانات
2.فرز وتبويب البيانات والتحضير للخوارزميات
3.تصميم النموذج
4.تقييم النتائج
5.اختيار ادوات التنقيب عن البيانات للكشف عن الغش والاحتيال المحاسبي

5.المحاسبة القضائية: يمكن استخدام الكثير من ادوات النقيب عن البيانات في مجال المحاسبة القضائية لاجراء مجموعه متنوعة من التحليلات التي كثيرا ما تستخدم للكشف عن عمليات الغش في الشراء. فعلي سبيل المثال، يمكن استخدام ادوات التنقيب عن البيانات للكشف عن أنماط أوامر الشراء التي توضع فقط تحت حدود وضوابط الشراء، وكذلك فروق الأسعار للمنتجات المتماثلة والمتشابهة، بالاضافه إلى ذلك يمكن مقارنه كميات الطلبات الخاصة بالمخزون أو اللوازم بالكميات المطلوبة في الفترات السابقة وتحليلها من قبل البائعين، وتقييمها مقارنه بالزيادات أو النقصان في المبيعات، ويمكن أيضا النظر في اتجاهات الأسعار، بما في ذلك المقارنات التاريخية الحالية والفروق السعرية العالية الانخفاض علي مدي فترات فصليه أو سنوية.

6.ترشيد قرار اختيار المشاريع الاستثمارية: يمكن استخدام ادوات التنقيب عن البيانات في ترشيد القرارات ومنها قرار اختيار المشاريع الاستثمارية وذلك للاسباب الاتية:

أ. ان اعداد دراسات الجدوى الاقتصادية لا يعد كافياً لاختيار المشاريع التي يمكن تنفيذها وذلك لوجود معايير اخرى تستوجب اخذها بنظر العناية عن اختيار تلك المشاريع من قبل الجهات التشريعية او المسؤولة عن اتخاذ اختيار المشاريع الاستثمارية.

ب. ظهور الكثير من العوامل الانية لم تكن موجودة سابقا اصبحت عوامل مؤثرة في اختيار المشاريع منها عوامل بيئية واجتماعية فضلاً عن العوامل السياسية والاقتصادية

ت. ان اجتياز المشاريع لاختبارات الموازنة الرأسمالية يجعلها جميعها مجدية للاستثمار الا ان عملية الاختيار بين تلك المشاريع يجعل عملية الاختيار عملية صعبة جداً لوجود متغيرات عديدة.

ث. وجود الكثير من البدائل تصلح ان تكون مشاريع استثمارية فضلاً عن وجود العديد من المشاريع التي يتطلب تنفيذها والتي تزداد باستمرار نتيجة النمو السكاني

المبحث الثاني / الاطار المفاهيمي للشبكات العصبية

اولاً: تعريف الشبكات العصبية

عرفت الشبكات العصبية بانها تتألف من مجموعة من العقد تؤدي نوع خاص من الحساب بشكل جماعي وان كل عقدة هي وحدة قياسية حسابية صغيرة وان هذه العقد يمكن إن تعمل بشكل متوازي حيث أنها تعتمد على تفاعلات فيما بينها او كيفية ارتباطها وهناك بعض العلماء عرفها على أنها :

- نماذج رياضية تحاكي صفات الأنظمة البايولوجية التي تعالج المعلومات بطريقة متوازية مؤلفة من عناصر بسيطة نسبيا تدعى عصيبات .

- هي صنف أو كيان بسيط من الخوارزميات الرياضية التي تصاغ بشكل مخططات(Graphs) تصنف هذه المخططات إلى عدد كبير من الخوارزميات وهذه الخوارزميات تقدم الحلول لعدد من المشاكل المعقدة (Humphreys, Zumac 1992.382)

كما عرفتها (الزبيدي والدليمي، 2008 ، 4) الشبكات العصبية على انها العلم الذي يهتم بدراسة الأساليب الرياضية التي يمكن صياغتها بالاعتماد على المحاكاة للخلايا البايولوجية في الكائنات، اذ تتميز الخلايا العصبية بالسرعة العالية في معالجة البيانات كما تتميز بقدرتها على التعلم والتعامل مع أنماط مختلفة من البيانات.

وعرفها (Minin, 2006 ,2) نموذج لمعالجه المعلومات المستوحاه من طريقة عمل نظم الخلايا العصبية البيولوجية في الدماغ ، والعنصر الرئيسي في هذا النموذج هو الهيكل الجديد لنظام تجهيز المعلومات. وهي تتألف من عدد كبير من عناصر المعالجة الشديدة الترابط (الخلايا العصبية) التي تعمل في انسجام لحل مشاكل محدده.

ويرى الباحث انها محاولة لايجاد تقنيات حاسوبية تصمم لمحاكاة الطريقة التي يؤدي بها الدماغ البشري مهمة معينة، وتتشابه معه في أنها تكتسب المعرفة بالتدريب المستمر وخزن هذه المعرفة ومن خلال ذلك يتم الوصول الى الحل الامثل للمسالة قيد الدراسة اذ تكتسب الشبكة المعرفة بالتجربة من خلال الاطلاع على البيانات التاريخية والاستفادة من التجارب السابقة.

ثانياً: انواع الشبكات العصبية

هناك عدة انواع من الشبكات العصبية وهي كما يلي: (ابراهيم ،2015، 56):

1.الشبكات العصبية ذات التغذية الامامية:(FFNN) Feed Forward Neural Networks

هي الشبكة العصبية الأكثر استخداما"، وتتألف من عدد من طبقات الخلايا العصبية التي يطلق عليها الوحدات والتي يتم ترتيبها في شكل طبقي وهناك اتجاه واحد فقط لحركة تدفق المعلومات تبدا من طبقة المدخلات والطبقة المخفية وطبقة المخرجات ولاتوجد هناك تغذية عكسية او دورات (Sutskever ،2013،8)

وللشبكات العصبية ذات التغذية الامامية (FFNN) عدة مزايا هي كما يلي (Bataineh، 2012,26):
أ. قدرة التنبؤ في أي مدخلات أو استقراء التدريب خارج الشبكة، بعد تدريب الشبكة ستكون قادرة على التنبؤ بأي مدخلات جديدة ، حتى تلك خارج حدود التدريب
ب. يعمل بشكل جيد للعديد من التطبيقات ، وخاصة تركيب منحنى السلسلة الزمنية البيانات (أي البيانات التي تأتي في أوقات وقيم مختلفة).

2.الشبكات العصبية ذات التغذية المرتجعة (Feed Back Neural Networks)

هي نوع من الشبكات العصبية التي تأخذ مدخلات متسلسلة وتنتج مخرجات متسلسلة من خلال مشاركة البيانات بين الخطوات الزمنية، وإلتي ادت الى نتائج مذهلة في معالجة اللغات الطبيعية والتسميات التوضيحية للصور، وقد تكون الشبكة ذات التغذية المرتجعة مباشرة اي تقوم بتكرار ناتج العصبون الى مدخلاته، او قد تكون الشبكة ذات التغذية المرتجعة المجاورة ، وتشمل توصيل مخرجات الخلايا العصبية الى المدخلات المجاورة، اي انها الشبكات التي تجد لمخرجاتها طريقا" خلفيا" مرة اخرى لتصبح مدخلات لكي تعطي افضل النتائج الممكنة (ابوعابدة،2015، 48)

3.الشبكات العصبية ذات الترابط الذاتي (Auto Associative Neural Networks)

هي الشبكات التي تلعب كافة العناصر المكونة لها دورا" نموذجيا"، يتمثل في استقبال المدخلات وبث المخرجات في الوقت نفسه (عابد،2015، 44). ويتم اعداد الشبكة العصبية بأشكال مختلفة حيث يتم اتصال الخلايا العصبية بطرائق مختلفة مما يعطي أشكالا عديدة للشبكة، وعند معالجة المعلومات تقوم كل عناصر المعالجة الحسابية المستندة اليها في الوقت نفسه بطريقة المعالجة المتوازية لمحاكاة طريقة عمل المخ البشري، وتتضمن بعض الأمثلة الهيكلية للشبكات العصبية (بوزيدي وعيشوش،48،2017) :

1.شبكة مكونة من طبقتين ذات تغذية في الاتجاه الأمامي .
2.شبكة مكونة من طبقتين ذات تغذية في الاتجاه الامامي والعكسي.
3.شبكة مكونة من طبقة واحدة ذات تغذية عكسية جانبية .
4. شبكة مكونة من عدة طبقات.

حدد (Haykin, S., 2009) انواع أساليب التعليم في الشبكات العصبية بالقول ان بالإمكان استخدام طريقتين لتدريب الشبكات العصبية كوسيلة للتعلم او المحاكاة بين النظام والمستخدم وهذه الطرائق هي :

1- Supervised (التعليم بوجود مشرف).
2- Unsupervised(التعليم مع عدم وجود مشرف).

1.الطريقة الأولى :

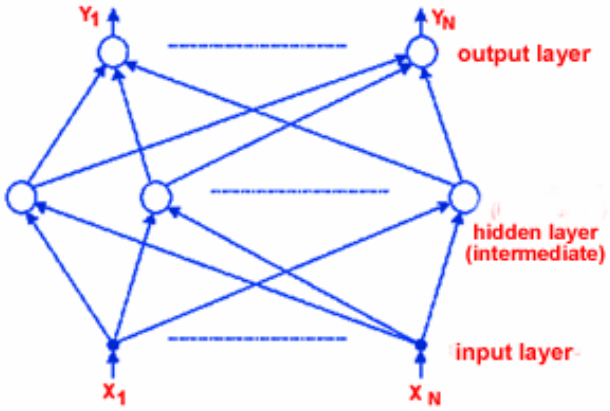
يفترض وجود مشرف خلال عملية التدريب لكل نموذج ويستخدم للاختبار ويتم خلال (target output) وتتضمن نموذج المدجلات إضافة إلى الإخراج المرغوب لتحديد (real output) وان في هذه العملية يتم إجراء مقارنة بين الإخراج الحقيقي إلى ان يتم تنقيص أداء الشبكة بعد إعطاء مصفوفة الأوزان ويحدد متجه الخطأ (real vector) حيث انه يقارن بين (output vector) مع متجه المخرجات والمدخلات للشبكة ويقارن أيضا مع الإخراج المرغوب لتحديد الأخطاء من خلال المعادلة الاتية :

Real output- target output = error

2.الطريقة الثانية:

وفي هذه الطريقة لايوجد مشرف لتقديم النماذج المطلوبة لذا يجب على النظام إن يتعلم بواسطة الاستكشاف والكيفية للخواص أو العوامل المهيكلة في نموذج الإدخال للمستخدم التجريبي ويجب إن يتم هذا التعلم من خلال تقوية الأوزان المنتخبة للعقد لكي تتطابق النماذج التجريبية ويميز التعلم بهذه الطريقة بصوره متكررة إلى إن يتم استقرار الأوزان .

وتعد شبكة (Multilayer Perceptron) هي الأكثر شهره والأكثر استخداما من انواع الشبكة العصبية. وفي معظم الحالات تنقل الإشارات داخل الشبكة في اتجاه واحد من المدخلات إلى المخرجات ليس هناك حلقه وإخراج كل الخلايا العصبية لا تؤثر في الخلايا العصبية نفسها



شكل (1)

معمارية شبكة Multilayer Perceptron

المصدر:

Marius, other, 2009, Multilayer Perceptron and Neural Networks, University of Craiova, Issue 7, Volume 8, p579

ثالثاً: خصائص واهمية الشبكات العصبية

تتميز الشبكات العصبية بالعديد من الخصائص أهمها (الصالح واخرون ، 2009 ، 9)

1. تعتمد على أساس رياضى قوى.
2. تمثل إحدى تطبيقات تكنولوجيا التشغيل الذاتى للمعلومات التى تقوم على محاكاة العقل البشرى.
3. تقبل أى نوع من البيانات الكمية أو النوعية.
4. لها القدرة على تخزين المعرفة المكتسبة من خلال الحالات التى يتم تشغيلها على الشبكة.
5. يمكن تطبيقها فى العديد من المجالات العلمية المختلفة

وقد اضافت (نزيه ، 2004 ، 5) الخصائص الاتية:

1. القدرة على اشتقاق المعنى من البيانات المعقدة أو الغير دقيقة
2. القدرة على التعلم كيفية القيام بمهام الاعتماد على البيانات بواسطة التدريب أو التجربة الأولية
3. بإمكانها خلق تنظيمها الخاص، وتمثيل البيانات التي تستلمها أثناء عملية التعلم
4. حسابات الشبكات العصبونية الاصطناعية قد تنفذ بشكل متوازي

واكد (Awodele ,Jegede,2009 , 12) على ان اهمية الشبكات العصبية تتضح من خلال قدرتها على اشتقاق معنى جديد من حجم كبير من البيانات المعقدة أو غير الدقيقة، اذ يمكن استخدامها لاستخراج أنماط والكشف عن الاتجاهات التي هي معقده للغاية للوصول الى نتائج لم تلاحظ من قبل البشر أو غيرها من تقنيات الكمبيوتر. ويمكن التفكير في شبكه العصبية المدربة بوصفها "خبيراً " في فئة المعلومات التي أعطيت لتحليلها ثم يمكن استخدام هذا الخبير لتقديم إسقاطات لحالات جديده من الاهتمام والاجابه عن الاسئلة "ماذا لو" فضلاً عن اهميتها بسبب المزايا الأخرى مثل التعلم التلقائي، والتنظيم الذاتي، وانجاز العملية في الوقت الحقيقي.

ويرى (Kingma ،2010,19) الدافع وراء الاهتمام بالشبكة العصبية هو لكونها بسيطة ومترابطة للغاية، وتمثل الوظائف الرياضية المعقدة بشكل فعال ، فضلا على انها مرنة جدا في التصميم ، ويمكن أن تكون الشبكة منفصلة تلقائيا فيما يتعلق بالحالات الداخلية، والمدخلات أو الأوزان من خلال خوارزمية backpropagation ، مما يجعل من السهل تنفيذ التحسين عن طريق الانحدار التدرج.

Anns

ويرى الباحث ان اهمية الشبكات العصبية تاتي من قدرتها على التعلم لنموذج العلاقات غير الخطية والمعقدة ، وهو أمر مهم حقا لأنه في الحياة الحقيقية هناك العديد من العلاقات بين المدخلات والمخرجات غير الخطيه والمعقده أيضا. كما ان الشبكات العصبية يمكنها التعميم بعد التعلم من المدخلات الاوليه وعلاقاتها اذ يمكن ان نستنتج العلاقات غير المرئية بين البيانات مما يمكن من تعميم النموذج على البيانات، فضلاً على انه وخلافا للعديد من تقنيات التنبؤ الأخرى لا تفرض الشبكات العصبية إيه قيود علي متغيرات المدخلات (مثل كيفيه توزيعها) مما يجعلها أفضل نموذج يمكن ان يستخدم مع البيانات ذات التقلبات العالية والتباين غير الثابت، نظرا لقدرته علي تعلم العلاقات الخفية في البيانات دون فرض اي علاقات ثابته في البيانات وهذا أمر مفيد جدا في التنبؤ بالسلاسل الزمنيه المالية (مثل أسعار الأسهم) حيث تكون تقلبات البيانات عاليه جدا

رابعا : مجالات استخدام الشبكات العصبية في المحاسبة والتدقيق

تعددت الدراسات المتعلقة بتطبيقات الشبكات العصبية في المحاسبة والتدقيق للافادة من مزاياها المتعددة فضلاً عن ازمة الثقة التي حدثت بعد انهيار بعض الشركات الكبرى مثل انرون اذ ظهرت مطالبات لاستخدام الشبكات العصبية في المحاسبة والتدقيق ولعل اهم مجالات استخداماتها هي: (جمعة ،2012، 206):

1.توجد اهمية لاستخدام الشبكات العصبية في مجالات التنبوء بالعوائد المتوقعة وتحليل الفرضية المحاسبية الخاصة بالاستمرارية واكتشاف الغش والخطأ كما ان انها تحسن من كفاءة عملية التدقيق بشأن اداء اختبارات المراجعة غير الضرورية لشركات لا تحتوي بياناتها على غش وهذا يؤدي الى تحسين كفاءة وفعالية احتبارات التدقيق.

2.ان الشبكات العصبية اكثر قدرة من النماذج الاحصائية التقليدية في النبوء بفشل الشركات وفي التنبوء بالشركات التي تقاريرها المالية بها تلاعب

3.ان استخدام الشبكات العصبية في التنبوء بارصدة الحسابات التي كانت جيدة كما ان قدرة الشبكات العصبية على اكتشاف الاخطاء المهمة الناتجة عن المعاملات الوهمية.

ويرى الباحث ان اهم المجالات التي يمكن استخدام الشبكات العصبية في المحاسبة وتدقيق الحسابات هي:

1.ان استخدام الشبكات العصبية في مجالات التنبوء يكون فاعلاً على وجه العموم لذلك يمكن استخدامها بفعالية في اعداد الموازنة لوحدات القطاع الخاص اذ تبتدء بالتنبوء بكمية المبيعات. فضلاً عن امكانية استخدامها في التنبوء في اعداد تقديرات الموازنة العامة للدولة من خلال التنبوء بقيمة النفقات للعام القادم

2.يمكن استخدام الشبكات العصبية في كل عمليات التقدير التي تعتمد في المحاسبة ومنها تقدير المخصصات مثل مخصص الديون المشكوك في تحصيلها ...

3. يمكن استخدام الشبكات العصبية في ترشيد العديد من القرارات ومنها ترشيد عملية الاختيار بين المشروعات الاستثمارية.

4.يمكن استخدام الشبكات العصبية في تحديد مخاطر التدقيق بكل انواعها

5.يمكن استخدام الشبكات العصبية في الكشف عن حالات الاحتيال والغش

6.يمكن استخدام الشبكات العصبية في مجال تحليل الحساسية لنقطة التعادل سواء لمنتج واحد او عدة منتجات

7.يمكن استخدام الشبكات العصبية في تحديد القيمة السوقية للموجودات النادرة

المبحث الثالث / استخدام التنقيب المحاسبي عن البيانات في تصميم برنامج لترشيد قرار المفاضلة بين المشاريع الاستثمارية باستخدام الشبكات العصبية (حالة دراسية)

لغرض بيان كيفية استخدام التنقيب عن البيانات في ترشيد قرارالاختيار بين المشاريع الاستثمارية ومثل هكذا قرار يمكن ان يتخذ من جهات متعدد مثلاً مجالس المحافظات او هيئات الاستثمار او مجلس النواب سوف يتم استخدام احد ادوات التنقيب عن البيانات وهي الشبكات العصبية في عملية اختيار المشاريع الاستثمارية ونفترض الحالة الدراسية الاتية:

ان المشاريع الاستثمارية الآتية تجاوزت مرحلة اعداد الدراسات الاقتصادية والفنية بكفاءة لغرض إجازتها ومن الطبيعي ان تكون هذه المشاريع الاستثمارية في مجالات متعددة فمثلاً منها الزراعية والصناعية والخدمية واذا كانت قد تجاوزت مرحلة اعداد دراسات الجدوى الاقتصادية فكيف سيتم الاختيار من قبل الجهات التي تمتلك قرار اختيار المشاريع الاستثمارية مثل البرلمان او مجالس المحافظات.

وقد توافرت البيانات الآتية باستخدام نظام المعلومات المحاسبية حول هذه المشاريع وكان لابد من اتخاذ قرار باختيار أحد هذه المشاريع الاستثمارية في حدود الموارد المالية المتاحة، وكما هو واضح في الجدول (1)

الجدول (1) المشاريع الاستثمارية المقترحة التي تجاوزت دراسة الجدوى الاقتصادية

المشاريع	نسبة الربح إلى رأس المال	عدد القوى العاملة التي سيتم توظيفها	جلب العملة الصعبة سنوياً	درجة التلوث البيئية والمسؤولية الاجتماعية	مبلغ الدعم الضريبي السنوي	مبلغ دعم المنتجات المحلية السنوي	استغلال الموارد المحلية	إحلال الواردات	تلبية الحاجات الأساسية	الترابط مع الصناعات القائمة	المعيار الأمني (غسيل الأموال)	الكفاءة الإدارية
1	10%	20	100,000	10	1,000,000	750,000	20	20	70	30	10	15
2	20%	25	250,000	20	1,500,000	800,000	45	15	90	40	20	15
3	25%	30	80,000	8	700,000	1,000,000	75	30	15	50	25	30
4	15%	25	70,000	11	600,000	1,100,000	15	40	25	30	15	40
5	20%	35	110,000	19	900,000	740,000	21	31	84	30	10	50
6	10%	36	65,000	18	1,000,000	810,000	44	26	89	80	5	55
7	25%	37	70,000	6	800,000	910,000	43	28	54	84	1	56
8	30%	40	170,000	20	1,100,000	1,000,000	41	34	27	55	15	58
9	15%	25	190,000	4	950,000	1,200,000	10	37	29	65	22	59
10	10%	28	200,000	15	850,000	1,300,000	31	44	62	67	18	57
11	20%	29	80,000	14	1,000,000	900,000	32	56	55	59	19	64
12	30%	25	65,000	18	780,000	860,000	35	57	49	68	17	62
13	27%	26	70,000	16	890,000	840,000	36	61	46	16	22	63
14	26%	27	80,000	17	900,000	890,000	60	62	36	34	24	61
15	24%	25	100,000	13	1,200,000	1,000,000	65	63	47	59	26	61
16	19%	20	95,000	10	990,000	960,000	66	64	49	68	25	56
17	18%	26	90,000	11	900,000	1,000,000	54	71	48	35	23	58
18	14%	31	100,000	14	1,000,000	980,000	56	45	49	16	32	59
19	15%	30	76,000	9	850,000	860,000	61	55	55	49	31	56
20	16%	26	69,000	8	900,000	870,000	71	61	75	57	34	53
21	22%	24	71,000	11	1,100,000	990,000	51	62	85	59	33	56
22	23%	25	72,000	10	780,000	1,000,000	53	37	69	64	36	57
23	24%	26	71,000	9	840,000	990,000	43	39	43	67	38	58
24	20%	22	81,000	12	910,000	900,000	61	33	45	51	33	51
25	21%	23	75,000	13	840,000	870,000	55	34	46	52	34	53
المجموع	70%	100	500,000	49	3,800,000	3,650,000	155	105	200	150	100	100

المصدر: الجدول من اعداد الباحث

ولغرض تطبيق النموذج الذي تم اختياره يجب توحيد نموذج تقييم المعايير بصيغة ارقام موحدة واعداد الجدول وفق صيغة الشبكة العصبية (Multilayer Perceptron) اذ تم تحويل البيانات الى صيغة موحدة وهي صيغة النسبة بقسم الجزء على الكل وكانت المدخلات في الشبكة العصبية وكما يأتي:

الجدول (2)

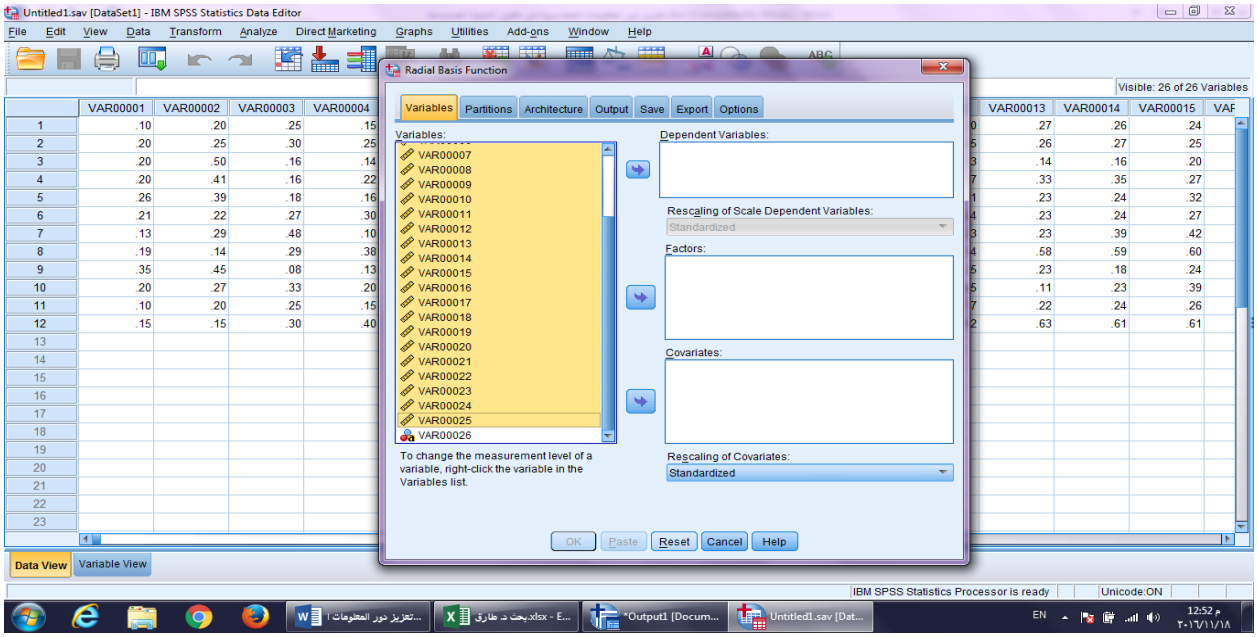
اعداد البيانات بصيغة شبكة
(Multilayer Perceptron)

المشاريع	نسبة الربح إلى رأس المال	عدد القوى العاملة التي سيتم توظيفها	جلب العملة الصعبة سنوياً	درجة التلوث البيئية والمسؤولية الاجتماعية	مبلغ الدعم الضريبي السنوي	مبلغ دعم المنتجات المحلية السنوي	استغلال الموارد المحلية	إحلال الواردات	تلبية الحاجات الأساسية	الترابط مع الصناعات القائمة	المعيار الأمني (غسيل الأموال)	الكفاءة الإدارية
1	0.1	0.2	0.20	0.20	0.26	0.21	0.13	0.19	0.35	0.20	0.10	0.15
2	0.2	0.25	0.50	0.41	0.39	0.22	0.29	0.14	0.45	0.27	0.20	0.15
3	0.25	0.3	0.16	0.16	0.18	0.27	0.48	0.29	0.08	0.33	0.25	0.30
4	0.15	0.25	0.14	0.22	0.16	0.30	0.10	0.38	0.13	0.20	0.15	0.40
5	0.2	0.35	0.22	0.39	0.24	0.20	0.14	0.30	0.42	0.20	0.10	0.50
6	0.1	0.36	0.13	0.37	0.26	0.22	0.28	0.25	0.45	0.53	0.05	0.55
7	0.25	0.37	0.14	0.12	0.21	0.25	0.28	0.27	0.27	0.56	0.01	0.56
8	0.3	0.4	0.34	0.41	0.29	0.27	0.26	0.32	0.14	0.37	0.15	0.58
9	0.15	0.25	0.38	0.08	0.25	0.33	0.06	0.35	0.15	0.43	0.22	0.59
10	0.1	0.28	0.40	0.31	0.22	0.36	0.20	0.42	0.31	0.45	0.18	0.57
11	0.2	0.29	0.16	0.29	0.26	0.25	0.21	0.53	0.28	0.39	0.19	0.64
12	0.3	0.25	0.13	0.37	0.21	0.24	0.23	0.54	0.25	0.45	0.17	0.62
13	0.27	0.26	0.14	0.33	0.23	0.23	0.23	0.58	0.23	0.11	0.22	0.63
14	0.26	0.27	0.16	0.35	0.24	0.24	0.39	0.59	0.18	0.23	0.24	0.61
15	0.24	0.25	0.20	0.27	0.32	0.27	0.42	0.60	0.24	0.39	0.26	0.61
16	0.19	0.2	0.19	0.20	0.26	0.26	0.43	0.61	0.25	0.45	0.25	0.56
17	0.18	0.26	0.18	0.22	0.24	0.27	0.35	0.68	0.24	0.23	0.23	0.58
18	0.14	0.31	0.20	0.29	0.26	0.27	0.36	0.43	0.25	0.11	0.32	0.59
19	0.15	0.3	0.15	0.18	0.22	0.24	0.39	0.52	0.28	0.33	0.31	0.56
20	0.16	0.26	0.14	0.16	0.24	0.24	0.46	0.58	0.38	0.38	0.34	0.53
21	0.22	0.24	0.14	0.22	0.29	0.27	0.33	0.59	0.43	0.39	0.33	0.56
22	0.23	0.25	0.14	0.20	0.21	0.27	0.34	0.35	0.35	0.43	0.36	0.57
23	0.24	0.26	0.14	0.18	0.22	0.27	0.28	0.37	0.22	0.45	0.38	0.58
24	20	0.22	0.16	0.24	0.24	0.25	0.39	0.31	0.23	0.34	0.33	0.51
25	21	0.23	0.15	0.27	0.22	0.24	0.35	0.32	0.23	0.35	0.34	0.53

المصدر: الجدول من اعداد الباحث

وبعد ذلك تم ادخال الجدول على برنامج (SPSS) وتم استخدام نموذج Neural Network ضمن البرنامج وكما هو واضح في الشكل (2)

استخدام النموذج تم استخراج مجموع النقط التفضيلية بكل مشروع وكما يلي:



شكل (2)

استخدام شبكة (Multilayer Perceptron) ضمن برنامج (SPSS)

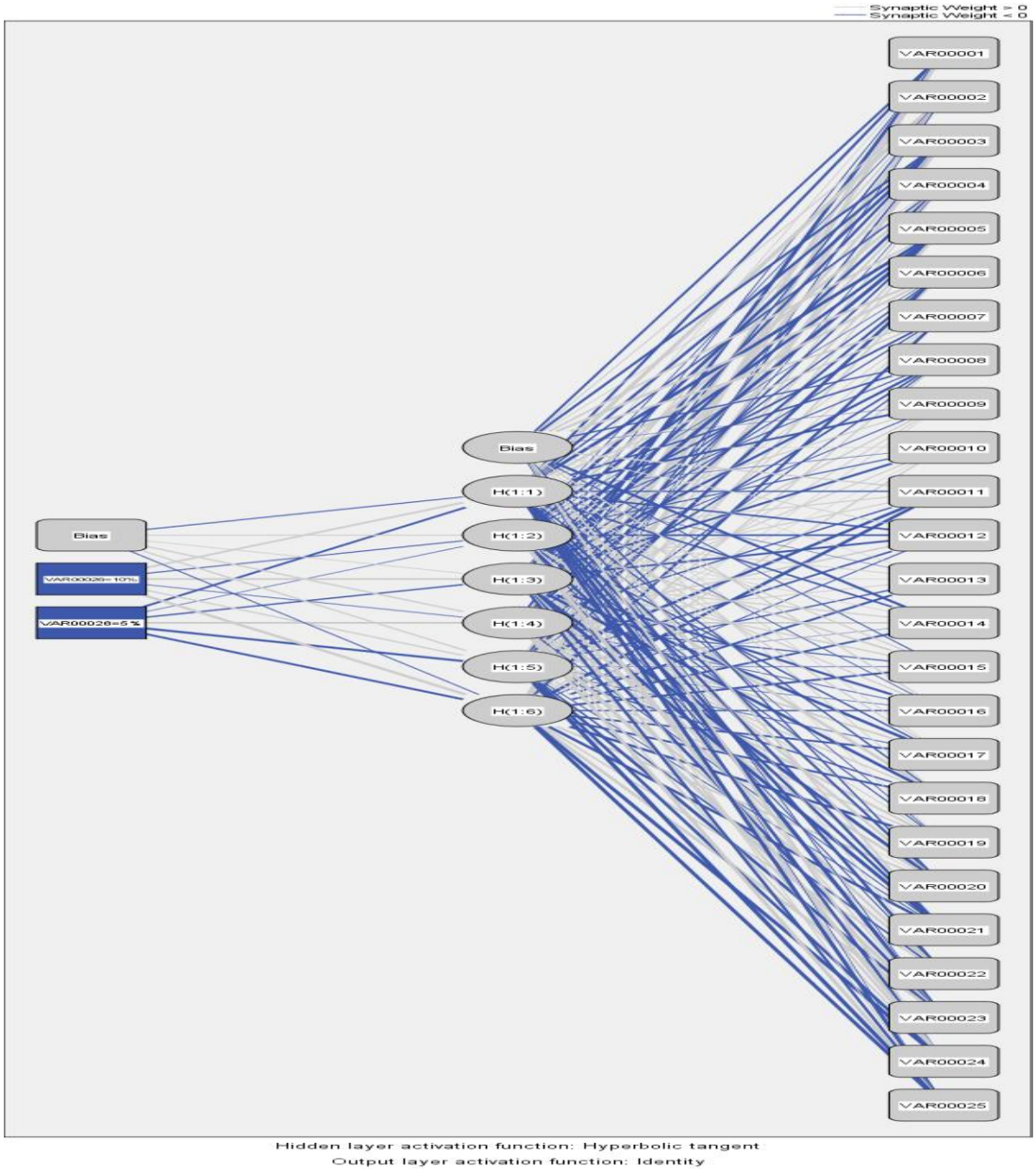
وبعد ذلك تم تنفيذ التحليل وكانت المخرجات كما يلي:

Case Processing Summary

		N	Percent
Sample	Training	8	66.7%
	Testing	4	33.3%
Valid		12	100.0%
Excluded		0	
Total		12	

ويتضح ان نسبة التدريب قد بلغت 66.7% بينما كانت نسبة الاختبار 33.3% وهي مؤشرات لكيفية تنفيذ الشبكة

شكل(3)
طبقات شبكة (Multilayer Perceptron) لعملية الاختيار بين المشاريع



جدول (3)
عملية التدريب والاختيار بين المشاريع

Model Summary

Training		Sum of Squares Error	70.207
		Average Overall Relative Error	0.802
	Relative Error for Scale Dependents	VAR00001	0.801
		VAR00002	**0.437**
		VAR00003	1.034
		VAR00004	0.781
		VAR00005	0.973
		VAR00006	1.109
		VAR00007	1.064
		VAR00008	1.185
		VAR00009	0.88
		VAR00010	0.973
		VAR00011	0.695
		VAR00012	0.805
		VAR00013	0.577
		VAR00014	0.724
		VAR00015	0.89
		VAR00016	0.748
		VAR00017	0.567
		VAR00018	0.502
		VAR00019	0.643
		VAR00020	0.825
		VAR00021	0.497
		VAR00022	1.07
		VAR00023	0.659
		VAR00024	0.813
		VAR00025	0.807
		Stopping Rule Used	consecutive 1
		Training Time	00:00.0
Testing		Sum of Squares Error	33811.31
		Average Overall Relative Error	1.308
	Relative Error for Scale Dependents	**VAR00001**	**0.974**
		VAR00002	1.007
		VAR00003	1.051
		VAR00004	1.395
		VAR00005	1.203
		VAR00006	1.23
		VAR00007	1.756
		VAR00008	0.991
		VAR00009	2.825
		VAR00010	1.961
		VAR00011	18.908
		VAR00012	14.373
		VAR00013	15.651
		VAR00014	4.654
		VAR00015	35.018
		VAR00016	20.641
		VAR00017	14.581
		VAR00018	6.273
		VAR00019	5.111
		VAR00020	2.549
		VAR00021	3.055
		VAR00022	5.324
		VAR00023	24.055
		VAR00024	1.307
		VAR00025	1.307

وكان المشروع المختار حسب تنفيذ الشبكة هو المشروع رقم (1) بعد ان اجرت الشبكة عملية تدريب لاتخاذ القرار ثم مرحلة الاختبار، اذ ان المؤشرات كانت تشير في مرحلة التدريب الى مشروع اخر ولكن عند اكتمال مرحلة الاختبار تم اختيار المشروع (1) ويتضح ذلك في جدول model summary اذ يتم اختيار المشروع لاقل نسبة خطأ تم تدريب الشبكة عليه

ومن خلال العرض السابق يمكن القول انه تم اثبات فرضيات البحث اذ تم اعتماد الشبكات العصبية وهي احدى ادوات التنقيب المحاسبي عن البيانات المحاسبية وتم من خلاله اختيار المشروع الاستثماري الامثل وهو قد ادى الى توفير اساس علمي مقبول من كل الاطراف لعملية اختيار المشاريع استثمارية فضلاً عن اختصار الوقت والجهد المبذول في اتخاذ هكذا قرار كما انه تم اثبات ان اعتماد التنقيب المحاسبي عن البيانات مكننا من اعتماد أساليب الذكاء الصناعي في معالجة تلك البيانات لتوفير معلومات مفيدة لترشيد القرارات الاستثمارية.

الاستنتاجات والتوصيات

توصل البحث الى مجموعة من الاستنتاجات وهي:

1. ان التزايد المضطرد بحجم البيانات عموماً ولاسيما البيانات المحاسبية يجعل من التعامل وفق الاطر التقليدية مسالة صعبة جداً كما انه يؤدي الى ضياع الوقت والجهد في استخلاص المعلومات.
2. ان ظهور كثير من المتغيرات الانية نتيجة المتغيرات في بيئة العمل يتطلب وجود ادوات تقنية لديها المرونة الكافية للتعامل معها.
3. ان ادوات التنقيب عن البيانات لديها القدرة في استنتاج علاقات بناء على قواعد البيانات الموجودة لديها لم تكن موجودة في السابق.
4. ان ادوات التنقيب عن البيانات لديها القدرة في التنبوء بالمستقبل لمتغير تتوفر لديها عنه بيانات ماضية.
5. ان التنقيب عن البيانات يمكن من التعامل مع حجم بيانات كبير جداً بسهولة ويسر وبكلفة ووقت قليل جداً.
6. المزايا التي تتمع بها ادوات التنقيب عن البيانات تمكنها في التعامل مع حجم البيانات المضطرد للبيانات المحاسبية
7. المجالات التي يمكن ان يستخدم بها التنقيب عن البيانات المحاسبية هي التنبوء بإفلاس الشركات، وتقدير مخاطر الائتمان، والإبلاغ المستمر، والضائقة المالية، والتنبؤ بأداء الشركات.
8. يمكن ان يستخدم التنقيب عن البيانات المحاسبية في مجال التدقيق في الاجراءات التحليلية وكذلك اكتشاف الغش والاحتيال وتحديد مخاطر التدقيق
9. يمكن ان يستخدم التنقيب عن البيانات المحاسبية في مجال المحاسبة القضائية
10. يمكن ان يستخدم بها التنقيب عن البيانات المحاسبية في مجال ترشيد قرار اختيار المشاريع الاستثمارية

ومن خلال العرض السابق للاستنتاجات التي تم التوصل اليها يمكن تقديم التوصيات الاتية:

1. اعتماد النموذج الذي قدمه البحث وهو شبكة (Multilayer Perceptron) وتوجد هذه الشبكة ضمن برنامج (SPSS) الذي يتيح امكانية استخدام هذه الشبكة بسهولة في ترشيد قرار اختيار المشاريع المنفذة في مجالس المحافظات
2. توجيه الدراسات نحو اعداد بحوث ورسائل واطاريح لدراسة امكانية اعتماد التنقيب عن البيانات في مجالات المحاسبة وتدقيق الحسابات للافادة من المزايا المتعددة التي توفرها ادوات التنقيب عن البيانات المحاسبية
3. توجيه الجهات المهنية المحاسبية والتدقيقية في العراق لاعتماد ادوات التنقيب المحاسبي عن البيانات في توفير معلومات مفيدة في ترشيد القرارات

المصادر

اولاً: المصادر باللغة العربية

1.ابراهيم، خالد عبد الغفور صالح،2015، دراسة المتغيرات المؤثرة في ارباح التمويل الاصغر باستخدام التحليل العاملي والشبكات العصبية، اطروحة دكتوراه، كلية الدراسات العليا ، جامعة السودان للعلوم والتكنولوجيا.
2.ابو عابدة، ألفت فتحي سالم، 2015، استخدام الطرق الإحصائية في التنبؤ بأسعار الذهب العالمية، رسالة ماجستير، كلية الاقتصاد والعلوم الادارية، جامعة الأزهر،غزة.
3.الدوري، زكريا مطلك، داليا عبد الحسين احمد، (2007) دور تنقيب البيانات في تطوير اداء المنظمة ، مجلة العلوم الاقتصادية ،مجلد 13 عدد 48 ، جامعة بغداد
4.بوزيدي وعيشوش ، 2017 ،
5.جمعة، أحمد حلمي ،2012، إستخدام الشبكات العصبية الإصطناعية في إكتشاف الأخطاء الجوهرية في البيانات المالية، المؤتمر العلمي السنوي الحادي عشر ذكاء الأعمال واقتصاد المعرفة، جامعة الزيتونة الأردنية، كلية الاقتصاد والعلوم الإدارية ، الأردن
6.حسين، احمد حسين (2004) نظم المعلومات المحاسبية الاطار الفكري والنظم التطبيقية ، الاسكندرية ، الدار الجامعية.
7.عابد، ياسر عبد الله،2015، دراسة مقارنة بين الأساليب الإحصائية لدراسة العوامل المؤثرة على تعدد الزوجات في الأراضي الفلسطينية، رسالة ماجستير، كلية الاقتصاد والعلوم الادارية، جامعة الأزهر، غزة.
8.عباس محسن البكري، غيث اسماعيل(2005) ، تصميم منظومة لتمييز أنماط الإشكال الهندسية باستخدام الشبكات العصبية ،جامعة بابل
9.العلي،عبد الستار العلي، عامر إبراهيم قنديلجي، غسان العمري (2011) المدخل إلى إدارة المعرفة، دار المسيرة للنشر والتوزيع والطباعة، الطبعة 3
10. الصالح، فروم محمد، بوجعادة إلياس، سليماني عز الدين ، 2009، دور أنظمة المعلومات المعتمدة على الذكاء الاصطناعي في عملية صنع القرارات الإدارية ، الملتقى الوطني السادس حول دور التقنيات الكمية في اتخاذ القرارات الإدارية ، الجزائر
11. الزبيدي، لهيب محمد، حنان حامد الدليمي (2008) استخدام الشبكة العصبية الاصطناعية نيوكوكنترون في تمييز الارقام العربية المكتوبة ، المؤتمر الاول ، جامعة الموصل
12. مها عبد الله محمد البدراني، (2008) ، استخدام شبكة القاعدة الشعاعية في تشخيص امراض الاطفال، المجلة العراقية للعلوم الاحصائية
13. نزيه، منى (2004) ، نماذج السلاسل الزمنية والشبكات العصبية في التنبؤ بأسعار الأسهم في بورصة الأوراق المالية"، معهد الدارسات والبحوث الإحصائية ، جامعة القاهرة.

ثانياً: المصادر باللغة الانكليزية

1.Bataineh، Mohammad Hindi،2012، Artificial Neural Network For Studying Human Performance, Thesis، University of Iowa.
2.Calderon, T.G. & J.J. Cheh,2002 "A Roadmap for Future Neural Networks Research in Auditing and Risk Assessment", International Journal of Accounting Information Systems, Volume 3, Issue 4
3.Crows, Two و 1999, Corporation: Introduction to Data Mining and Knowledge Discovery, Third Edition,

4.Efraim, Turban, & Dorothy, Leidner,2008, Information Technology for Management, 6th edition, john wiley & sons, New Jersy
5.Haykin, S. (2009), Neural Networks and Learning Machines, 3rd ed., Pearson Education, Upper Saddle River, NJ.
6.Humphreys, R, & Zumach,W. ,1992, Automated star/galaxy discrimination with neural networks. The Astromical Journal, Vol., 1, No., 103
7.Koh, H.C. , 2004و "Going Concern Prediction using Data Mining Techniques", Managerial Auditing Journal, Volume 19, No 3,
8.Kingma، 2010، Improving Score Matching for learning statistical models of natural images، Thesis ،Universities Utrecht.
9.Lin T.Y., 2007, Attribute (Feature) Completion The Theory of Attributes from Data Mining Prospect , , San José State University
10. Marius,Constantin Popescu, Valentina E. Balas, Liliana Perescu-Popescu, Nikos Mastorakis, 2009, Multilayer Perceptron and Neural Networks, University of Craiova1, WSEAS TRANSACTIONS on CIRCUITS and SYSTEMS Issue 7, Volume 8, July
11. Masood Zamani , 2017 ,Protein Secondary Structure Prediction Evaluation and a Novel Transition Site Model with New Encoding Schemes , The University of Guelph , Canada,
12. Minin و Alexey, 2006و The Neural -Network Analysis Data Filters,
13. Oludele Awodele , Olawale Jegede,2009,Neural Networks and Its Application in Engineering ,Proceedings of Informing Science & IT Education Conference (InSITE) , Babcock University, Nigeria
14. Sutskever، Ilya،2013، Training Recurrent Neural Networks، Doctor of Philosophy، University of Toronto.

Accounting Mining Data Using Neural Networks (Case study)

ABSTRACT

Business organizations have faced many challenges in recent times, most important of which is information technology, because it is widely spread and easy to use. Its use has led to an increase in the amount of data that business organizations deal with an unprecedented manner. The amount of data available through the internet is a problem that many parties seek to find solutions for. Why is it available there in this huge amount randomly? Many expectations have revealed that in 2017, there will be devices connected to the internet estimated at three times the population of the Earth, and in 2015 more than one and a half billion gigabytes of data was transferred every minute globally. Thus, the so-called data mining emerged as a technique that aims at extracting knowledge from huge amounts of data, based on mathematical algorithms, which are the basis for data mining. They are derived from many sciences such as statistics, mathematics, logic, learning science, artificial intelligence, expert systems, form-recognition science, and other sciences, which are considered smart and non-traditional.

The problem of the research states that the steady increase in the amount of data, as well as the emergence of many current areas that require different data due to the contemporary environment of business organizations today, make information systems unable to meet the needs of these current organizations, and this applies exactly to accounting information systems as they are the main system in business organizations today. These systems have been designed to meet specific needs that make it impossible today to meet the different needs according to the contemporary environment of business organizations today, as well as failing to deal with the amount of data generated by the information technologies.

The research proposes two main hypotheses. First, the adoption of accounting data mining leads to providing data that the accounting information system was unable to provide before, as well as to shortening the time and effort required to obtain it. Second, the adoption of accounting exploration of data enables the adoption of artificial intelligence methods in processing such data to provide useful information to rationalize decisions.

The research leads to a number of conclusions, including that the steady increase in the amount of data in general, and the accounting data in particular, makes dealing with traditional frameworks a very difficult issue and leads to loss of time and effort during extracting information. In addition, the emergence of many current variables as a result of changes in the work environment requires the presence of technical tools, which have enough flexibility to deal with them. Moreover, data mining tools have the ability to derive relationships based on their existing databases that were not available before.

The research presents a number of recommendations, most important of which is the need to adopt the model presented by the research, i.e., Multilayer Perception, a network that exists within the (SPSS) program, which allows the possibility to use this network easily in rationalizing the decision to choose implemented projects in the provincial councils.

Key words: **Accounting Mining, Neural Networks, Multilayer Perceptron Network ، Training Neural Networks**