مجلة جامعة تشرين للبحوث والدراسات العلمية _ سلسلة العلوم الهندسية المجلد (31) العدد (1) 2009
Tishreen University Journal for Research and Scientific Studies - Engineering Sciences Series Vol. (31) No. (1) 2009

دراسة تأثير تغييرات بارامترات الشبكة العصبونية على أدائها

الدكتورة كندة أبو قاسم *
الدكتور بلال شيحا **
منذر سكيف ***

(تاريخ الإيداع 23 / 10 / 2008. قُبِل للنشر في 2009/2/1)

□ الملخّص □

تم في هذا البحث دراسة شبكة عصبونية ذات ثلاث طبقات وتصميمها، شبكة قادرة على تعلّم مجموعة معطيات كبيرة بمساعدة طريقة خطأ الانتشار العكسي، ودراسة تأثير تغيير البارامترات (خطوة التعلّم، عدد العقد، نوع تابع التفعيل من أجل عدد من إشارات الدخل المختلفة) والأثر الشديد الذي يسببه هذا التغيير في عمل الشبكة العصبونية، كما برهنت نتائج هذه التجارب على الحساسية الشديدة لاستجابة الشبكة العصبونية المصممة التي تعتمد على تقنية الانتشار العكسي لتغير هذه البارامترات.

الكلمات المفتاحية : شبكات عصبونية- خوارزمية الانتشار الخلفي- تعليم خوارزمية- طبقات.

* أستاذ مساعد - قسم هندسة الحاسبات والتحكم الآلي- كلية الهندسة الميكانيكية والكهربائية- جامعة تشرين-اللاذقية- سورية.
** مدرس- قسم هندسة الحاسبات والتحكم الآلي- كلية الهندسة الميكانيكية والكهربائية- جامعة تشرين-اللاذقية- سورية.
*** طالب دراسات عليا(ماجستير) - قسم هندسة الحاسبات والتحكم الآلي-كلية الهندسة الميكانيكية والكهربائية-جامعة تشرين-اللاذقية- سورية.

مجلة جامعة تشرين للبحوث والدراسات العلمية _ سلسلة العلوم الهندسية المجلد (31) العدد (1) 2009
Tishreen University Journal for Research and Scientific Studies - Engineering Sciences Series Vol. (31) No. (1) 2009

Studying the Effect of the Neural Network's Parameters Change on the Network Performance

Dr. Kinda Abo Khasem*
Dr. Belal Sheha**
Monzer Skef***

(Received 23 / 10 / 2008. Accepted 1 / 2 / 2009)

□ ABSTRACT □

This research aims at studying and designing a three-layer neural network, capable of dealing with massive data, relying on the back-propagation error method. It also explores the effect of parameters-change (learning step, number of nodes, type of activation function for a few different input signals) and the major impact, caused by this change on the neural network performance.

The results of these experiments have demonstrated the extreme sensitivity of the designed network which depends on the back-propagation mechanism to change these parameters.

Key words: neural networks, back propagation algorithm, learning algorithm, layers.

* **Associate Professor, Department of Computer and Automatic Control, Faculty of Mechanical and Electrical Engineering, Tishreen University, Lattakia, Syria.**
**** Assistant Professor, Department of Computer and Automatic Control, Faculty of Mechanical and Electrical Engineering, Tishreen University, Lattakia, Syria.**
***** Postgraduate Student, Department of Computer and Automatic Control, Faculty of Mechanical and Electrical Engineering, Tishreen University, Lattakia, Syria.**

مقدمة :

تطور حقل الشبكات العصبونية بناء على عوامل كثيرة مثل بيولوجيا الأعصاب ، الرياضيات ، علوم الكومبيوتر (الذكاء الصنعي) ، الفيزياء .....الخ .

بدأت النظرية الحديثة للشبكات العصبونية منذ أربعينيات القرن الماضي في أعمال Warren McCulloch و Walter Pitts اللذين برهنا على إمكانية قيام الشبكات العصبونية من حيث المبدأ بعملية حسابية أو منطقية أطلق عليها البنية الجديدة Perceptron. واستمر العمل في السبعينيات من قبل عدة باحثين إلى أن تم تطوير شبكة عصبونية قادرة على العمل بوصفها ذاكرة [1] . في هذه المقالة استخدمت خوارزمية الانتشار العكسي Back Propagation algorithm لتدريب شبكة عصبونية متعددة الطبقات من النوع Feed-Forward ، بحيث نأخذ ثلاث طبقات للشبكة العصبونية مع مدخل ومخرج وحيد . إن خوارزمية الانتشار العكسي هي الخوارزمية الأكثر شيوعاً الآن في عملية تعليم الشبكات العصبونية .

تعتمد فكرة هذه الخوارزمية على تعديل الأوزان بحيث يقل الخطأ بين قيمة الخرج الفعلي والقيمة المرغوب فيها للخرج[2] . وتتطلب هذه الطريقة حساب مشتق الخطأ بالنسبة إلى الأوزان؛ أي تتطلب معرفة كيفية تغير الخطأ بالنسبة إلى قيمته السابقة، وتعدل الأوزان بين الطبقة المخفية وطبقة الخرج بأن يضاف إلى الأوزان السابقة معدل تغير الخطأ في هذه الأوزان [3] تعاد هذه العملية طبقة طبقة حتى طبقة الدخل وهذا ما أعطى اسم الانتشار العكسي لهذه الخوارزمية .

أهمية البحث وأهدافه :

يتلخص هدف البحث في إبراز المشاكل التي تعاني منها الشبكات العصبونية الحالية ،ودراسة تأثير تغير البارامترات(خطوة التعليم ، عدد العقد ، نوع تابع التفعيل من أجل عدد من إشارات الدخل المختلفة)، وكيفية التغلب على الأخطاء الناتجة عنها بالاعتماد على خوارزمية الانتشار الخلفي. تكمن أهمية البحث في إيجاد البارامترات المثالية للشبكة العصبونية للحصول على أقل خطأ ممكن.

طريقة البحث ومواده :

تم في هذا البحث دراسة النقاط التالية:

- دراسة مرجعية في الشبكات العصبونية.
- اعتماد خوارزمية الانتشار العكسي.
- تأثير تغيير أهم البارامترات في عمل الشبكة العصبونية.
- مناقشة نتائج الدراسة وصياغة الاستنتاجات.

تصميم أداء الشبكة وتقييمها:

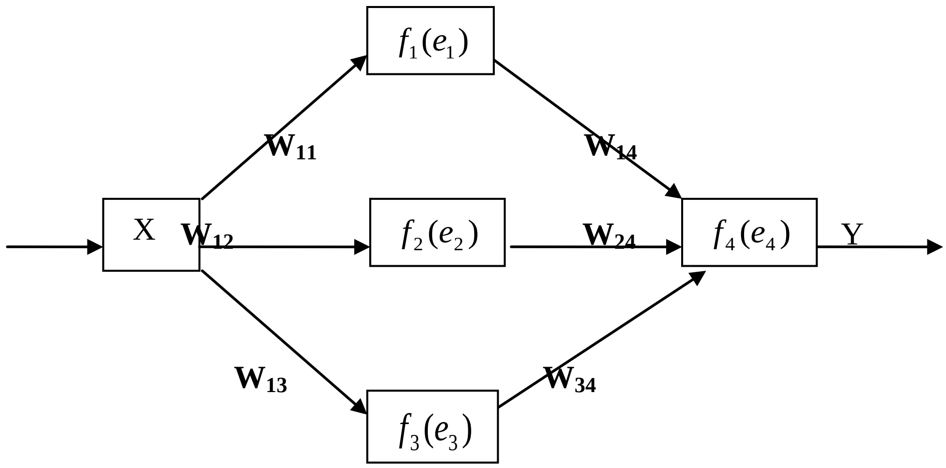
يتطلب تصميم أداء الشبكة وتقييمها الخطوات التالية :

- بناء شبكة عصبونية مؤلفة من ثلاث طبقات بطبقة مخفية وحيدة، وهي تكفي لتمثيل النظام [3].
- خوارزمية تدريب الشبكة العصبونية، وهي خوارزمية الانتشار العكسي BP .
- تابع لتقدير أداء الشبكة العصبونية، وهو تابع الخطأ(error).

الخوارزمية المستخدمة في البرمجة:

خوارزمية الانتشار العكسي

سنقوم بدراسة الشبكة العصبونية متعددة الطبقات موظفين خوارزمية الانتشار العكسي، من أجل توضيح هذه العملية نأخذ ثلاث طبقات للشبكة العصبونية مع مدخل ومخرج واحد كما هو موضح في الشكل (1)



الشكل(1) شبكة عصبونية متعددة الطبقات

كل عصبون مركب من وحدتين، الوحدة الأولى تنتج معاملات الأوزان وإشارات الدخل . الوحدة الثانية وهي التي تقوم باستدعاء تابع تنشيط العصبون، حيث يمثل $f(\ )$ تابع التنشيط الخاص بكل عصبون، كما يمثل x الدخل و y الخرج، و e تمثل مجموع جداء الدخل بالأوزان التابعة له، وتمثل w الأوزان التي تربط بين عصبونات الشبكة بعضها مع بعض المداخل والمخارج.

لدراسة الشبكة العصبونية نحن نحتاج لتدريب مجموعة معطيات $(X_1, X_2, ...)$. فعند كل عملية تدريب يتم تعديل قيم أوزان الشبكة العصبونية[4].

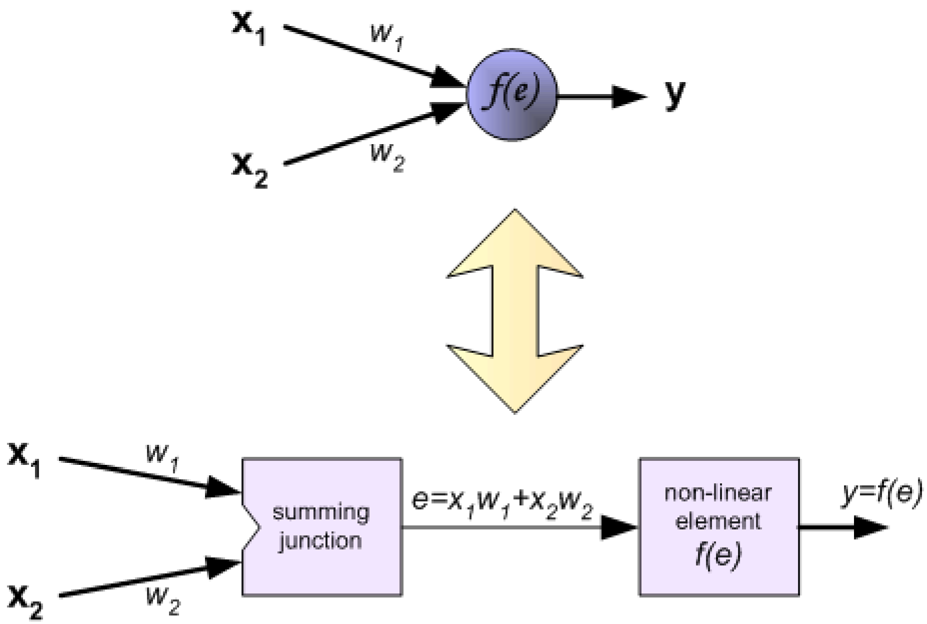
التعديل يحسب باستخدام الخوارزمية الموصوفة كما يلي :

كل خطوة تعليم تبدأ بتغذية إشارة الدخل من مجموعة التدريب، بعد هذه المرحلة نستطيع تحديد قيمة إشارات الخرج لكل عصبون في كل طبقة من الشبكة كما هو مبين في الشكل (2).

الانتشار العكسي(Back Propagation

هو خوارزمية للتعليم بواسطة مثال وفيه:

- في الشبكة متعددة الطبقات، تستطيع المعلومات التي تغذي عقد الدخل أن تتمثل داخلياً بحيث أن المخارج تتوالد من هذا التمثيل الداخلي [5].
- إن الخرج يتحدد من خلال العمليات الداخلية لعقد الشبكة.



الشكل(2) يبين كيفية حساب الخرج انطلاقاً من معطيات الدخل واعتماداً على بنية الشبكة العصبونية NN

سيناريو التعلم (The Learning Scenerio)

- الشبكة تقاد بواسطة مجموعة من الأمثلة من أزواج الدخل – الخرج (مجموعة التدريب) .
- الأوزان يمكن أن تعدل عند تقريب التابع من أزواج الدخل – الخرج المرسومة [4].
- بعدها تفحص الشبكة من خلال مقدرتها على أداء الدور المطلوب.
- تصحيح الخطأ يتم:

– في أثناء التدريب يوضع الدخل في الشبكة ويتدفق عبرها مولداً مجموعة من القيم على وحدات الخرج .

– الخرج الفعلي يتم حسابه ومن ثم يقارن مع الهدف المطلوب:

1- إذا وجد تطابق بين الخرج الفعلي والخرج الهدف فلا يحدث أي تغيير في الشبكة[2] .

2- وعند عدم وجود التطابق يتم التغيير.

موجز تاريخي عن الانتشار العكسي

- أول مقارب لهبوط الانحدار في الشبكات متعددة الطبقات كان في عام 1967م بوساطة Amari- الذي استخدم طبقة مخفية واحدة لتشكيل التصنيف اللاخطي.
- في عام 1974م اكتشف Werbos " ديناميكية التغذية العكسية ".
- في عام 1982م تكلم Parker على " تعلم المنطق" وظهر هذا ثانية في عام 1985م.

• في عام 1985م أصدر روميل هارت وهينتون وليم – التأثير الأعظم على انتشار الخوارزمية في مجموعة الشبكة العصبونية.

تعليم الخوارزمية(Learning Algorithm)

تعليم الخوارزمية يعتمد على هبوط منحنا الخطأ[6] ، إذ يعبر عن الخطأ كما يلي:

$$E = \sum_{P} E_P \qquad (1)$$

حيث E_P هي الخطأ من أجل دخل واحد، و E هي الخطأ الكلي.

$$E_P = \frac{1}{2}\sum_{i}\left(t_i - a_i\right)^2 \qquad (2)$$

حيث

a_i هو الخرج الفعلي.

i هو عدد التكرار.

سوف نجعل الأوزان تتغير بناء على التغيير في منحنا الخطأ

$$\Delta w = -\eta \nabla E \qquad (3)$$

حيث η هو عامل ثابت التدرج(ثابت التعليم) يحدد حجم الخطوة .

ويكون التغير في الوزن الذي يربط العصبون i مع العصبون j هو:

$$\nabla w_{ji} = -\eta \nabla_{ji} E = -\frac{\partial E}{\partial w_{ji}}$$

ويمكن التعبير عن الحد $\frac{\partial E}{\partial w_{ij}}$ اعتماداً على بنية الشبكة العصبونية NN [7].

$$\frac{\partial E}{\partial w_{ji}} = \frac{\partial E}{\partial a_j} \cdot \frac{\partial a_j}{\partial net_j} \cdot \frac{\partial net_j}{\partial w_{ji}} \qquad (4)$$

يتم حساب المشتق الجزئي الثالث في المعادلة (4) من خلال تحديد net_j.

حيث

$$net_j = \sum_{i=0}^{n} a_i * w_{ji}$$

$$\frac{\partial net_j}{\partial w_{ji}} = \frac{\partial}{\partial w_{ji}} \sum_{K=0}^{n} w_{jk} a_k = \sum_{k=0}^{n} \frac{\partial w_{jk} a_k}{\partial w_{ji}} \qquad (5)$$

بفحص المشتق الجزئي، نلاحظ بأن $\frac{\partial w_{jk}}{\partial w_{ji}}$ يساوي الصفر ما لم يكن $k = i$.

وبالتالي تصبح المعادلة (5):

$$\frac{\partial net_j}{\partial w_{ji}} = a_i$$

أما بالعودة إلى المشتق الجزئي : $\frac{\partial a_j}{\partial net_j}$ وبما أن $a_j = f(net_j)$

وباعتبار $f(net_j) = \frac{1}{1+e^{-net_j}}$

عندها نجد $\frac{\partial a_j}{\partial net_j} = f'(net_j) = \frac{d(1+e^{-x})^{-1}}{dnet_j}$

$$\frac{d(1+e^{-x})^{-1}}{dnet_j} = (-1)(1+e^{-x})^{-2}e^{-x}(-1)$$

$$= \frac{e^{-x}}{1+e^{-x}} \cdot \frac{1}{1+e^{-x}}$$

$$= \frac{1+e^{-x}-1}{1+e^{-x}} \cdot \frac{1}{1+e^{-x}} = \left(\frac{1+e^{-x}}{1+e^{-x}} - \frac{1}{1+e^{-x}}\right) \cdot \frac{1}{1+e^{-x}}$$

$$= (1-a_j)a_j$$

لذا يكون:

$$\frac{\partial a_j}{\partial net_j} = (1-a_j)a_j \qquad (6)$$

ولحساب المشتق الأول من المعادلة (4): $\frac{\partial E}{\partial a_j}$ نستفيد من المعادلة (2):

بالعودة إلى

$$E_P = \frac{1}{2}\sum_i (t_i - a_i)^2 \qquad (7)$$

حيث المجموع يكون على خرج وحدات الشبكة ، و يمكننا أن ندرس حالتين من أجل المشتق الجزئي

- J هو خرج وحدة
- J هو للطبقة المخفية(hidden layer)

إذا كانت j خرج وحدة ، عندها نستطيع أن نحسب المشتق ببساطة كما يلي:

$$\frac{\partial E}{\partial a_j} = \frac{\partial}{\partial a_j} \cdot \frac{1}{2}\sum_i (t_i - a_i)^2$$

$$= \sum_i (t_i - a_i)\frac{\partial (t_i - a_i)}{\partial a_j}$$

$$= (t_i - a_i)(-1) = -(t_i - a_i)$$

$$\partial_j = -(t_i - a_i)(1-a_j)a_j$$

أما إذا كانت a_j للطبقة المخفية(h)، فعندها نحن نحتاج إلى الاتكال على قانون التدرج المطبق على الوحدات k ، الوصل إلى الوحدة j [8].

$$\frac{\partial E}{\partial w_{ji}} = \delta \cdot w_{ij} \cdot f'(net_h) \cdot X_i \tag{8}$$

حيث :

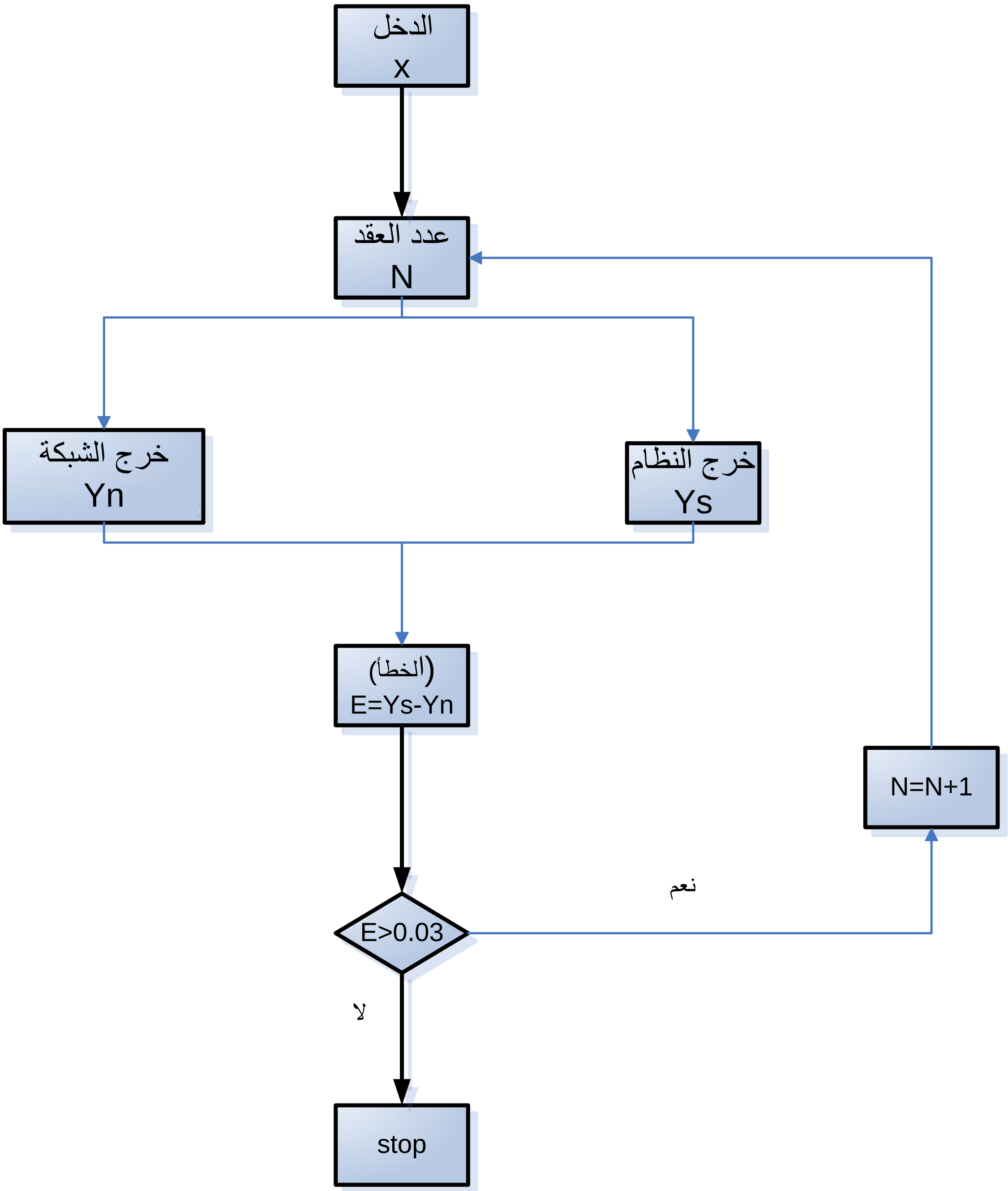
$$f'(net_h) = \frac{\partial out_h}{\partial net_h}$$

$$w_{ij} = \frac{\partial net_o}{\partial out_h}$$

$$x_i = \frac{\partial net_h}{\partial w_i}$$

الخوارزمية المعتمدة في الوصول للشبكة المناسبة (تصميم الشبكة):

بعد تصميم الشبكة تم تدريبها حيث حصلت أخطاء، فقمنا بمقارنة كل خطأ مع الخطأ الذي يسبقه مباشرة، فإذا تناقص الخطأ كان التدريب جيداً، وإذا لم يتناقص نقوم بزيادة عدد العقد ونستمر بهذه الزيادة حتى يبدأ تناقص الخطأ كما هو مبين في الشكل (3) .



الشكل (3):يبين الخوارزمية المعتمدة للوصول للشبكة المناسبة

النتائج والمناقشة:

نناقش في هذه الفقرة ثلاث نقاط أساسية وهي :

1- تغيير خطوة التعليم (STEP_W).

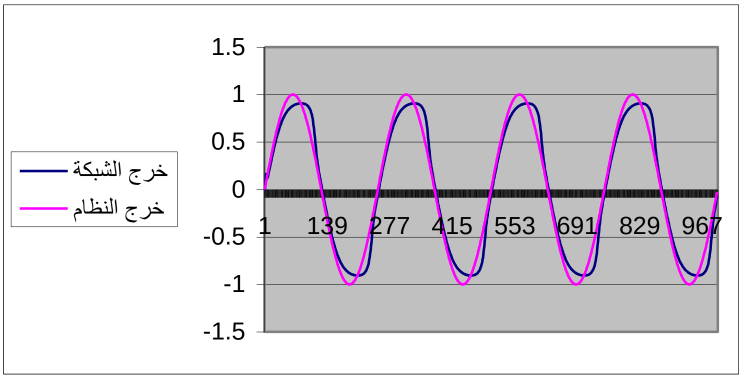
2- تغيير عدد العقد (NO_OF_NODE) . 3- إشارات دخل مختلفة.

أولاً- تغيير خطوة التعليم (STEP_W)

حيث أن تابع $\sin(2\pi t/250)$ هو التابع المستخدم بوصفه دخولاً على الشبكة.

الجدول(1): التأثيرات المختلفة لتغيير خطوة التعليم على أداء الشبكة العصبونية.

عدد مرات التنفيذ	عدد العقد الأعظمي	خطوة التعليم	مقدار الخطأ	عدد العقد بعد التنفيذ	مجموع مربع الخطأ (بعد التدريب)
1000	200	0.01	-0.002718	200	16.06
1000	200	0.001	-0.317749	200	188.83
1000	200	0.06	-0.000080	120	7.20

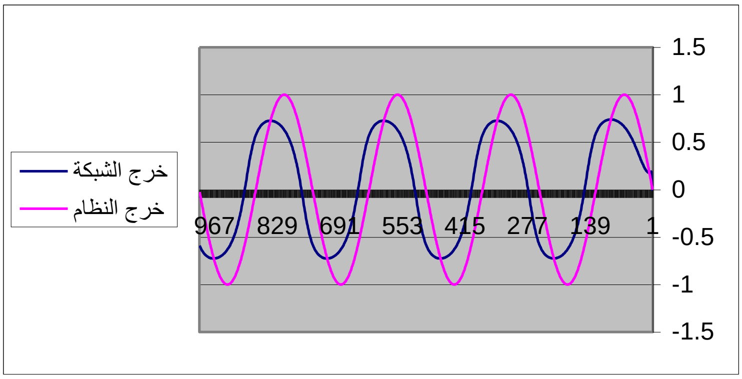


الشكل(4) يبين التقارب عندما خطوة التعليم = 0.01

عدد التكرار =1000.

عدد العقد الأعظمي=200.

خطوة التعليم =0.01.

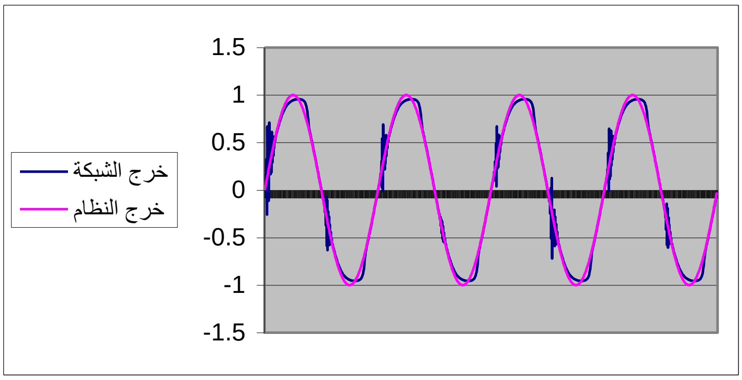


الشكل(5) تأثير خطوة التعليم عندما أصبحت=0.001

عدد التكرار =1000.

عدد العقد الأعظمي=200.

خطوة التعليم =0.001



الشكل(6) يبين التقارب عندما خطوة التعليم = 0.06

عدد التكرار=1000.

عدد العقد الأعظمي=200.

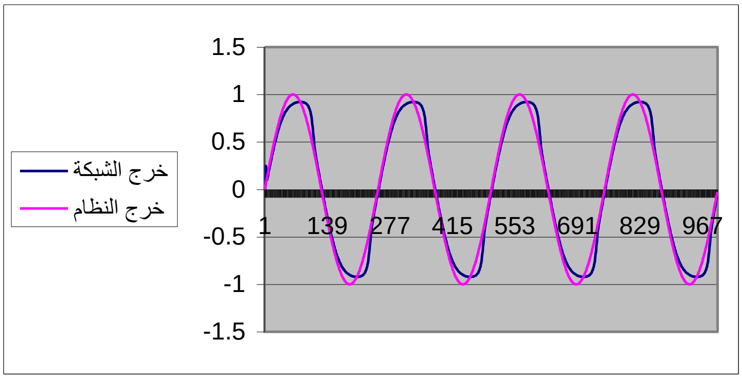
خطوة التعليم =0.06.

ثانيا – تغيير عدد العقد (NO_OF_NODE)

حيث إن تابع $\sin(2\pi t/250)$ هو التابع المستخدم بوصفه دخولاً على الشبكة

الجدول 2: التأثيرات المختلفة لتغيير عدد العقد على أداء الشبكة العصبونية.

عدد مرات التنفيذ	عدد العقد الأعظمي	خطوة التعلم	مقدار الخطأ	عدد العقد بعد التدريب	مجموع مربع الخطأ بعد التدريب
1000	50	0.06	-0.001157	50	10.39
1000	100	0.06	-0.000284	100	5.01
1000	150	0.06	-0.000125	150	3.29

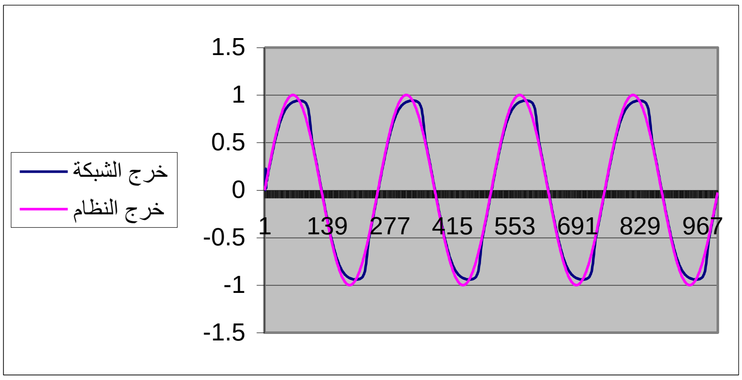


الشكل(7) يبين خرج الشبكة عندما عدد العقد =50

عدد التكرار=1000.

عدد العقد الأعظمي=50.

خطوة التعليم =0.06.

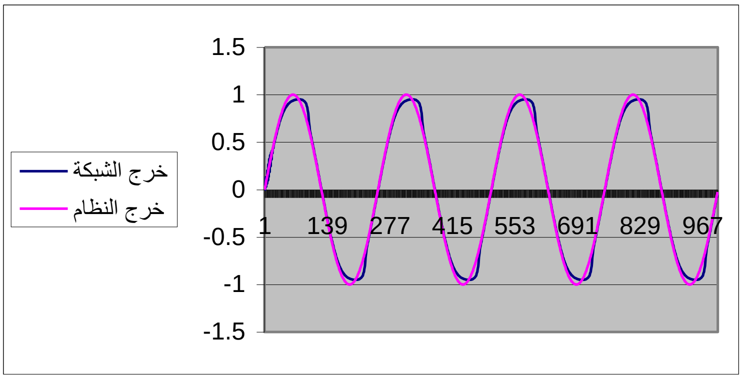


الشكل (8) يبين خرج الشبكة عندما عدد العقد =100

عدد التكرار =1000.

عدد العقد الأعظمي=100.

خطوة التعليم =0.06.



الشكل (9) يبين خرج الشبكة عندما كان عدد العقد =150

عدد التكرار =1000.

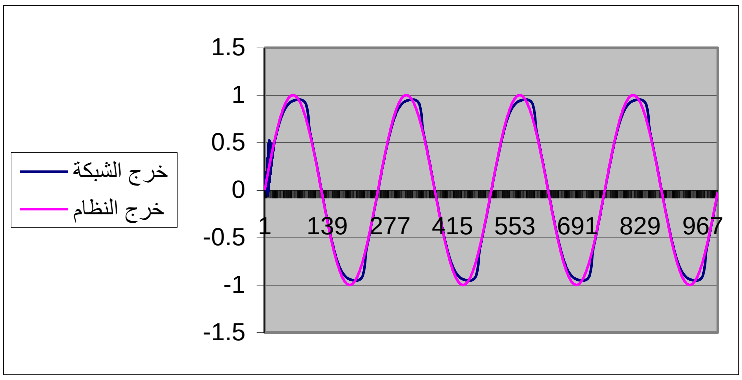
عدد العقد الأعظمي=150.

خطوة التعليم =0.06

ثالثا – إشارات دخل مختلفة :

الجدول 3: التأثيرات المختلفة لنوع إشارة الدخل على أداء الشبكة العصبونية.

نوع اشارة الدخل	عدد العقد الأعظمي	عدد التكرار	مجموع مربع الخطأ	مقدار الخطأ	عدد مرات التنفيذ	عدد العقد بعد التنفيذ
SIN	200	1000	3.16	-0.000102	1063	200
COS	200	1000	375.78	-0.003784	1063	200
SIN^2	150	2000	6.56	0.000001	2063	150
SIN*COS	200	1000	48.42	-0.306595	1063	115

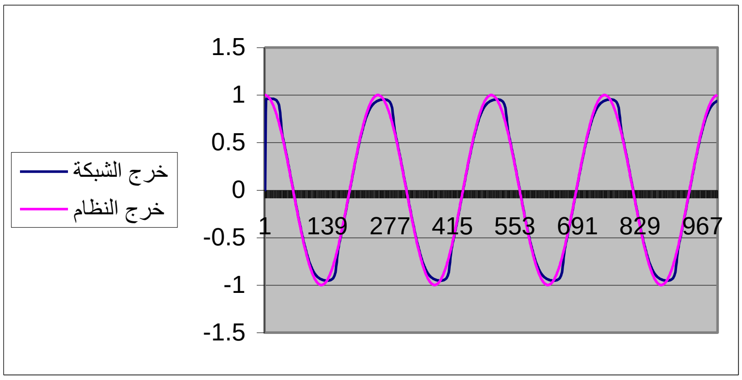


الشكل(10) يبين أداء الشبكة عندما كانت إشارة الدخل sin

عدد التكرار=1000.

عدد العقد الأعظمي=200.

خطوة التعليم =0.06

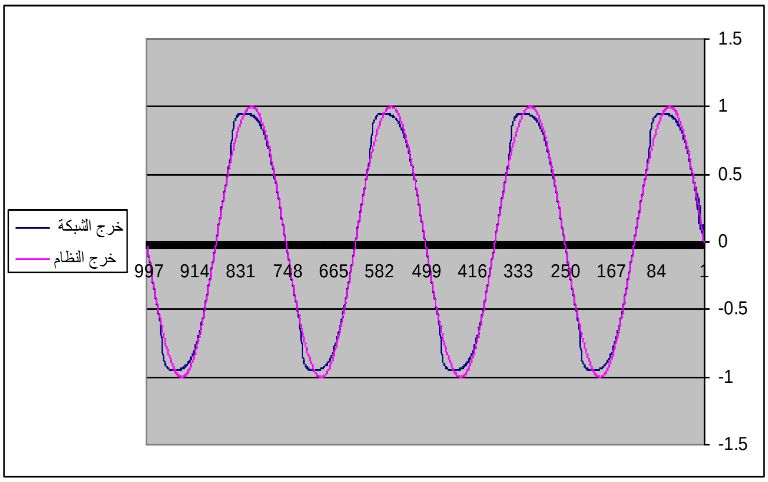


الشكل(11) يبين أداء الشبكة عندما كانت إشارة الدخل cos

عدد التكرار=1000.

عدد العقد الأعظمي=150.

خطوة التعليم = 0.05

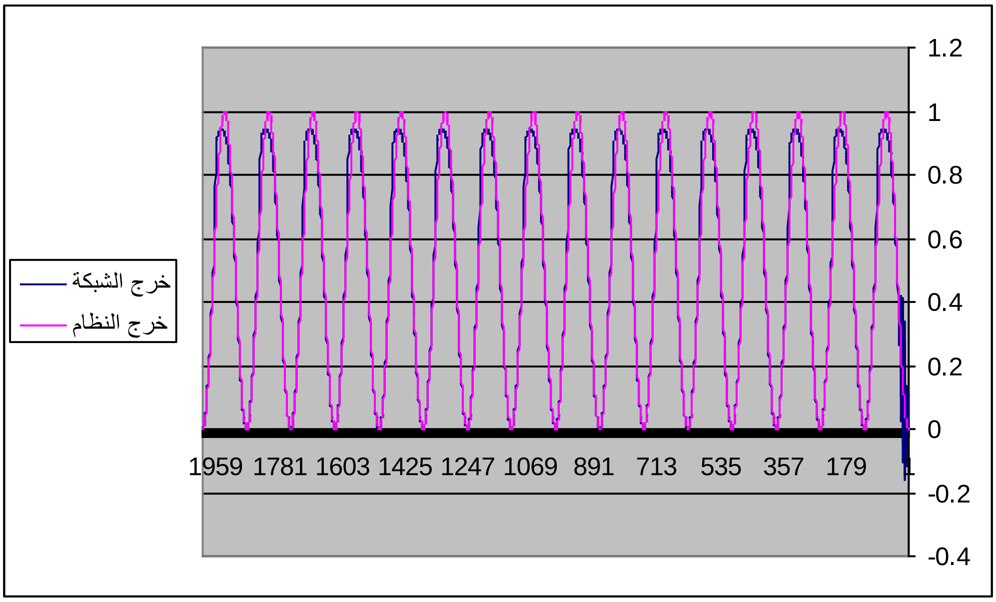


الشكل(12) يبين أداء الشبكة عندما كانت إشارة الدخل sin*cos

عدد التكرار =1000.

عدد العقد الأعظمي=200.

عدد العقد بعد التنفيذ = 115.



الشكل(13) يبين أداء الشبكة عندما كانت إشارة الدخل $\sin^2$

عدد التكرار =1000.

عدد العقد الأعظمي=150.

خطوة التعليم = 0.05

في الجدول (1) نجد أنه كلما نقصت خطوة التعليم (STEP_W) كلما أصبح التقارب بين الخرج الحقيقي والخرج المثالي أسوأ، وهذا مايوضحه الشكل(2) ؛ إذ عندما كانت خطوة التعليم تساوي 0.001 كان مجموع مربع الخطأ يساوي 188.83، والتباعد بين الخرجين أكبر ما يمكن .

أما في الجدول (2) فيبين أنه كلما زاد عدد العقد كلما كان التقارب أفضل ؛ إذ نجد في الشكل (9) أن عدد العقد (150) وهي أعلى قيمة في الجدول، وأن التقارب بين الخرجين هو الأفضل ، حيث قيمة مربع الخطأ (3.29) وهي أقل من باقي القيم لمربع الخطأ في الجدول (2) وبالتالي التقارب هو الأفضل عن هذه القيمة .

وفي الجدول (3) نلاحظ أنه على الرغم من إشارة الدخل مربع التابع الجيبي أكثر تعقيداً من سابقاتها، إلاّ أنه بعد أن تم اختيار البارامترات بشكل ملائم كان التقارب بين خرج النظام وخرج الشبكة جيداً كما هو مبين في الشكل (13).

الاستنتاجات والتوصيات :

1- بينت نتائج الدراسة أن خوارزمية الانتشار الخلفي تعدّ وسيلة جيدة لتدريب الشبكات العصبونية.

2- ينصح ببنية شبكة عصبونية تتألف من طبقة مخفية واحدة.

3- بينت الدراسة أن اختيار تابع السيغمويد (sigmoid) كان مناسباً؛ لأنه مكّن الشبكة من مقاربة تابع الهدف.

4- وقد بينت النتائج وجود عيوب باستخدام الخوارزمية المقترحة؛ لذلك ينصح باستخدام الخوارزمية الجينية للوصول إلى النتائج الأمثلية في تصميم الشبكات العصبونية.

المراجع :

1- ERIC,D.and PATRICK,N.", *Neural Networks"*,university of Manchester,Edition Eyrolles.Paris1991,145P.

2- SARTORI,A. and ANTSAKLIS,P.,"*Implementations of learning control systems using neural networks*,"IEEE control systems magazine,April 1992,57.

3- http://www.techguide.com.

4- LEE,J.,"*Astuy on speeding up learning of neural networks"*,The Institute of Electrical Engineering ,chung-Hua polytechnic institute,Taiwan,1994,105.

5- RUMELHART,D.,HINTON,G.and WILLIAMS,R.,"*learning internal Representations by error propagation,"*vol .1.MIT press,1986,536.

6- NG,S.,LEUNG,S. and LUK,A."*Ageneralized Back-propagation algorithm for faster convergence"*,IEEE international conference on neural networks,Washington D.C.,3-6,vol.l,1996,409-413.

7- ADNAN,S.,TAYFUN,M.&SINAN,U.,"*Determination of efficiency of flat-plate solar collectors using neural network approach*",Ankara,Turkey,2008,1533-1539.

8- MELLIT,A.,BENGHANEM,M.,&KALOGIROU,A.,*"Modeling and simulation of a stand-alone photovoltaic system using an adaptive artificial neural network : proposition for anew sizing procedure"*,Renewable Energy,2007,285-313.