مجلة القادسية لعلوم الحاسبات والرياضيات المجلد (1) العدد (1) السنة(2009)
المؤتمر العلمي الأول لكلية علوم الحاسبات والرياضيات المنعقد في 27-28 /أب / 2008

Page 202-208

دراسه تحليليه لتطور خوارزمية انسياب الخطأ خلفا العصبية من خلال عرض مشكلة الـ X-OR

وجدان عباس
المعهد الفني في النجف/قسم الحاسبات

الخلاصه

يهدف البحث الى توضيح مفهوم الشبكات العصبية وخصوصا شبكة الخطأ (Error Backpropagation BP) خلفا وطرح مشكلة X-OR وكيفية حلها وعمل محاكاة لكل نوع من انواع الخوارزميات في لغة Matlab وايجاد افضل خوارزمية من خلال مقارنة بين النتائج في حل المشكله. حيث تم استخدام معاملات لتطوير خوارزمية Standard BP ،بادخال قيمة MOMEUTUM والتي تحتفظ بجزء من التصحيح السابق لغرض تعميم عملية التعلم اما الخوارزمية الثانية فهي تعتمد على ADAPTIVE Learning Rate الذي يعطي مرونة اكبر للوصول الى الحل بوقت اقصر والخوارزمية الثالثة ERROR_DECAY انحلال الخطاء والتي تعطي نسبة تصحيح افضل.

1- مقدمة عامة (General IntroductioN)

ظهرت الحاسبات الرقمية في بدايتها لحل مسائل حسابية معقدة تستغرق وقتا طويلا لحلها يدويا، ولكن التطور الذي شهدتة الحاسبات و النجاح الذي حققته في تطبيقات متعددة جعلها اكثر من مجرد، اله حسابية (computing machine) الذي كان دافعا قويا للعلماء لتطوير واكتشاف عدة تقنيات، تحاول استغلال أمكانية الحاسبة لانجاز وظائف ذكية (intelligent tasks) مرتبطة بالسلوك ، البشري وتعد سهلة بالنسبة للبشر كمشكلة تميز الأنماط (pattern recognition)ومنها الأنظمة الذكية، (intelligent systems) وتشمل
(الأنظمة الخبيرة،والشبكات العصبية ،وانظمة المنطق المضبب، والخوارزميات الجينية) التي حاول بها الباحثون محاكاة قدرة الانسان على التعلم وادراك وتميز الظواهر، والأشياء المحيطة بة وقابلية على تطوير نفسه. نالت الشبكات العصبية (ARTIFICIAL NEURAL NETWORKS) اهتماما من قبل الباحثين كاحد موضوعات العصر المهمة التي حققت التحدي الذي يسعى الية باحثوا الذكاء الاصطناعي (ARTIFICIAL INTELLIGENCE) في تصميم انظمة ذات سلوك البشري في عملية اتخاذ القرار وحل المشاكل فظهرت الشبكات العصبية الاصطناعية لمحاكات الشبكات العصبية البيولوجية التي تعد بالاساس نماذجا رياضية (MATHEMATICAL MODELS) لمعالجة المعلومات. قدم روزنبلات(rosenblatt) عام 1958 اول تصميم لشبكة عصبية تتكون من مجموعة خلايا بالاعتماد على نموذج ماكلوج_بيتز [1] الا ان النموذج المقدم لم يستطع تخطي مجال المشاكل الخطيه. في علم 1986 قام روميلهارت ورفاقة (rumelhart et .al) بتصميم شبكة انسياب الخطأ خلفا EBPالتي تميزت بخوارزمية تعلم مشرف علية وتعتمد هذة الشبكة على المشتقة في عملية تعديل اوزانها بدلا من العتبة (threshold function)المستعملة في الشبكات الخطية وهذة الشبكة قادرة على حل مشاكل التقريب بتعقيدات مختلفة.

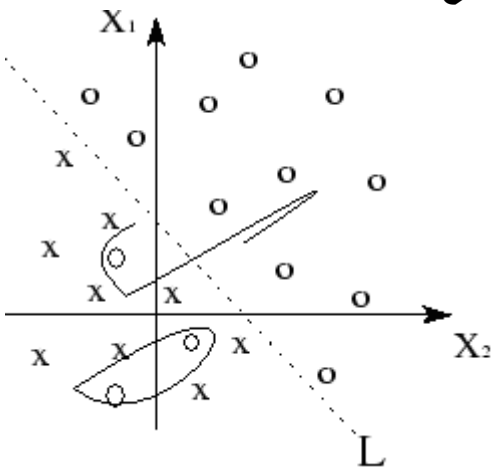
2- مشكلة X-OR problem

هي عبارة عن مشكلة اساسية تستخدم دائماللاختبار والمقارنة في الشبكات العصبية وX-OR هي دالة فصل غير خطية [3] [4][5] والجدول (1) يوضح دالة X-OR

جدول (1) مشكلة X-OR

A	B	A XOR B
1	1	0
0	1	1
1	0	1
0	0	0

لحل المشكلة بالترتيب نحتاج الى تقديم طبقه جديده الى الشبكه العصبيه ،هذه الطبقه تدعى عادة "الطبقه المخبأ" تسمح للشبكه خلق صيانه طبقات داخليه للمدخلات،هنا الشبكه مع الطبقه المخبأ سوف تنتج الجدول الحقيقي للX-OR والشكل (1) يوضح دالةX-OR



شكل (1) الحالة اللاخطيه لمشكلة X-OR

2-1-شبكه Xor:-

يتم إدخال البيانات إلى ألشبكه فإذا كانت المخرجات الناتجة لا تطابق المخرجات المرغوبه الشبكه تطلق اشاره خطأ(error signal) وتذهب تلك الاشاره الى الخلف خلال الشبكه وصولاً الى عقده الادخال كالخطأ المعبور ،تغير معماريه الشبكه لكي نقلل الخطأ لكي نقوم بذللك نعمل التالي :-

1:-نسمح لكل ارتباط ان يمتلك مجموعة اوزان عشوائية

2:- نضع" threshold" العتبه .

3:-في معماريه الشبكه تصبح العقدة فعاله اذا وفقط اذا الادخال الكلي كان اعظم من العتبه

لانملك العتبه وللحصول عليها يجب ان نغير داله فعاليه الشبكه ويكون تغير قانون الفعاليه كالتالي :-

التنشيط مساوي الى(1)اكثرمن(1-)ورفع الى قوى المجموع السالب

للادخال الكلي للعقده

حيث ان الإدخال الكلي هومجموع لكل المدخلات الى العقده المعطاه، الادخال يعيد الناتج الى العقده الفعاله التي ترسل المدخلات ووزن الاتصال ويوضح الجدول (1) جميع معاملات الشبكه في العمل الحالي.

الجدول (1) معاملات الشبكه التي تم استخدامها في هذا العمل

المعامل	القيمة
عدد عقد الادخال	2
عدد العقد في الطبقة المخبأه	2
عدد العقد في طبقة الاخراج	1
معامل التعلم	0.05
عدد نماذج التدريب	10
عدد نماذج الاختبار	4
الحد الأعلى للخطأ المسموح به	$5*10^{-5}$

3- خوارزمية انسياب الخطأ خلفا ERROR BACKPROPAGATION

تعتمد هذه الخوارزمية [1] على قيمة الخطأ الناتج في طبقة الاخراج ، حيث يتم احتساب الخطأ الكلي الناتج وثم إعادة توزيعه باتجاه الخلف ليؤثر على تعديل معاملات الاوزان لكل عقده وبشكل يتناسب مع مشتقة الخطأ لكل عقده. ويمكن تمثيله رياضيا بالمعادله (1)

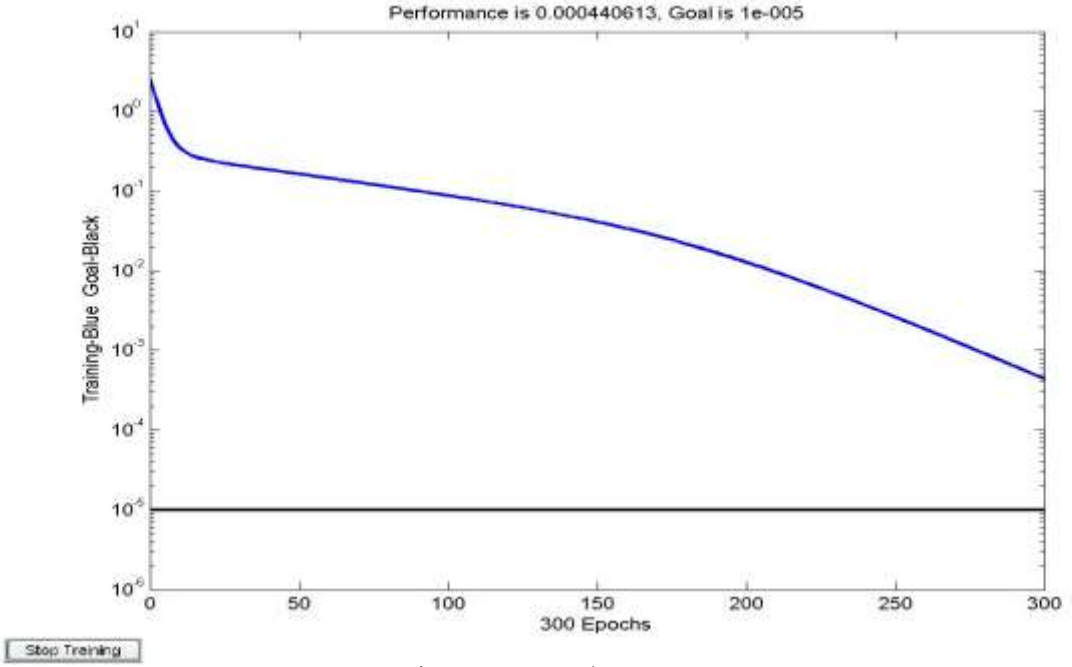
$$dX = lr * dperf/dX \qquad (1)$$

حيث ان dx تمثل نسبة التصحيح

Lr: عامل التعلم

dperf/dX: مشتقة الخطأ

يوضح الشكل (2) تدريب الشبكه وفق الخوارزميه القياسيه.



شكل (2) تدريب خوارزمية انسياب الخطأ خلفا القياسيه

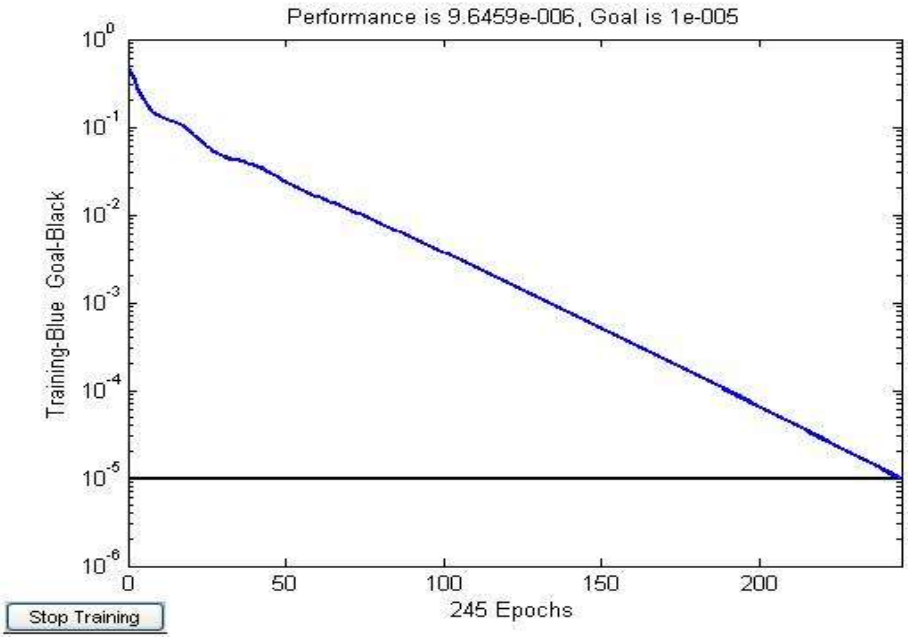
يوضح الشكل اعلاه تأخر الشبكه في الوصول الى الهدف المطلوب بالاضافه الى وقوعها في الخطأ المحلي اوالذي يوضحه الانحدار الحاد في بداية التدريب ومن ثم تراجع الشبكه في الوصول الى الهدف.

3-1-خوارزمية انسياب الخطأ خلفا باستخدام العزم Error Backpropagation With Momentum

وتعتمد هذه الخوارزمية على إدخال عامل العزم [6] ويمثل هذا العامل جزء من التعديل السابق للأوزان ، مما يساعد على تقليل نسبة تأرجح الشبكة حول local minima. يعتمد التعديل الحالي عادة على المشاهدة الحالية (قيمة الإدخال الحالي) لذلك يكون اتجاه التعديل مائلا وبشكل كبير الى شكل المدخل الحالي، لذلك فان اتجاه التعديل السابق سوف يقلل من شدة الخطأ في التعديل الحالي
ويتم تمثيلها رياضيا بالمعادلة (2)

$$dX = mc*dXprev + lr*(1-mc)*dperf/dX \qquad (2)$$

حيث ان mc يمثل معامل العزم
dXprev: نسبة التعديل السابقه
الشكل (3) يمثل تمرين الشبكه على 10 نماذج باستخدام عامل العزم



شكل(3) استخدام عامل العزم

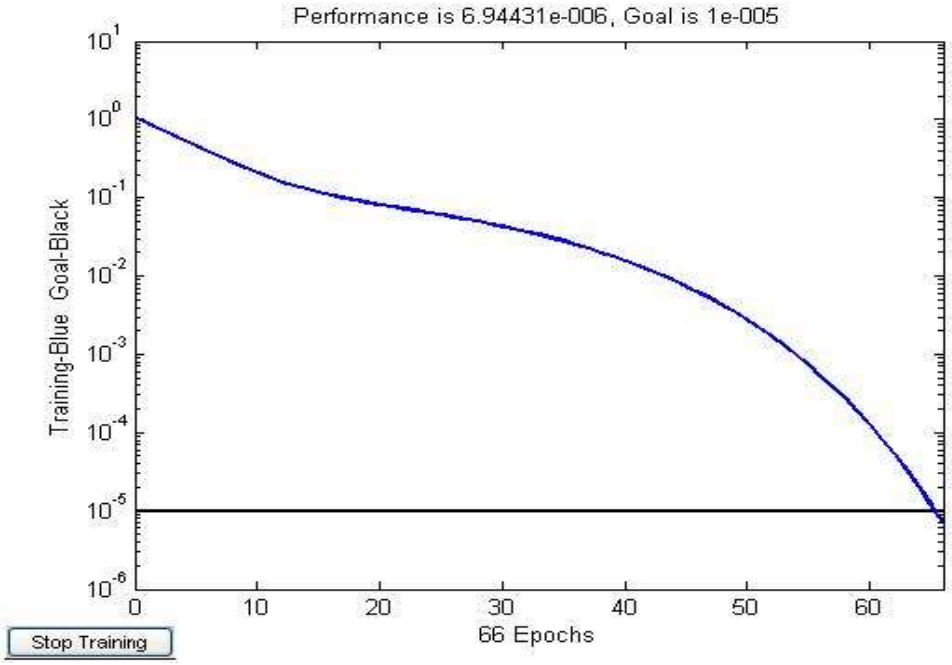
من الشكل (3) يتضح ان الشبكة استطاعت الوصول الى الهدف خلال الدوره 250 ، كما يبين الشكل ان الشبكه كانت بعيده عن مشكلة الوقوع في مشكلة الخطأ المحلي وهذا نتيجة تأثير عامل العزم في تنعيم انحدار الخطأ.

3-2- خوارزمية انسياب الخطأ خلفا باستخدام معامل تعلم قابل للتغير Error Backpropagation With Adaptive Learning Rate

تعاني شبكة انسياب الخطأ خلفا في بعض المسائل من وقوعها في منطقة هدف محلي ويجعلها عاجزة عن الوصول إلى الهدف الحقيقي بسبب ثبات عامل التعلم طول فترة التدريب.[7]
تعتمد هذه الخوارزمية على متابعة سلوك الشبكه في اثناء عملية التدريب، فتساعد في تعجيل زمن الوصول الى الهدف بوقت اقل وذلك باستخدام خطوات واسعه في منطقة الحل عندما تكون الشبكه بعيده عن الهدف واخرى دقيقه في مرحلة تقدم الشبكه نحو الهدف المطلوب.

$$dX = lr*dperf/dX \qquad (3)$$

يوضح الشكل (4) تمرين الشبكه باستخدام معامل التعلم القابل للتغير



الشكل (4) تمرين الشبكه باستخدام معامل التعلم القابل للتغير

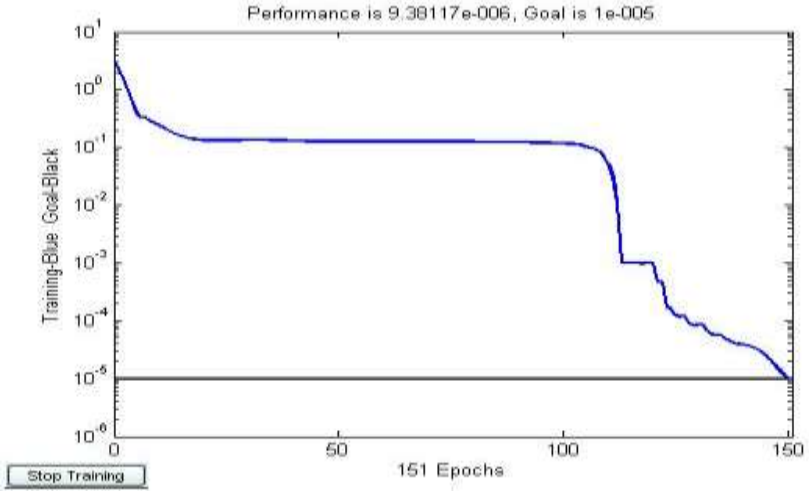
يوضح الشكل(4) الوصول السريع للشبكه الى منطقة الهدف مقارنة مع الشكل (3) . كما يعكس سلوك انحدار الخطأ والذي أخذ انحدارا بسيطا في بداية مرحلة التدريب ومن ثم تسارع الانحدار في مراحل التدريب المتقدمه والناتج عن تعديل معال التعلم كي يلائم منطقة الحل المحليه في كل مرحله.

3-3- خوارزمية انسياب الخطأ خلفا باستخدام العزم Error Backpropagation With Momentum And Adaptive Learning Rate

يمكن الاستفاده من خواص معامل التعلم المتغير وعامل العزم لتشكيل خوارزميه تعتمد على التغييرين كي تكون اكثر امانا في الوصول الى الهدف وذلك بالابتعاد عن مشكلة الخطأ المحلي والتأرجح حول منطقة الهدف. توضح المعادله (4) الصيغة الرياضيه لاستخدام كلا العاملين في خوارزمية انسياب الخطأ خلفا [9] [8].

$$dX = mc*dXprev + lr*mc*dperf/dX \qquad (4)$$

يوضح الشكل (5) تمرين الشبكه باستخدام معامل التعلم القابل للتغير وعامل العزم



الشكل (5) تمرين الشبكه باستخدام معامل التعلم القابل للتغير وعامل العزم

من خلال الشكل(5) نجد ان الشبكة استغرقت وقتا اطول في الوصول الى الهدف وذلك من خلال 150 دوره، الا ان الشبكه استطاعت التعلم بعد الدوره 100 وبشكل سريع.

4- الاستنتاجات Conclusion

من خلال استخدام النموذج القياسي لخوارزمية شبكة انسياب الخطأ خلفا نجد انها عرضة للوقوع في مشكلة الخطأ المحلي وكما موضح بالشكل(3)،حيث تصبح الشبكة ضمن منطقة خطأ مثاليه نسبيا(مقارنة مع المنطقه المحيطة). الا ان التعديل الناتج على هذه الخوارزميه اعطاها مرونه عاليه في القفز حول منطقة الخطأ المحلي من خلال استخدام قيمة العزم، والذي يقوم بتقليل اثر التعلم الخاطئ(تعديل الاوزان باتجاه لا يتناسب واتجاه الشبكة نحو اقل قيمة للخطأ) كما تم تجنب تأرجح الشبكه حول منطقة الهدف من خلال تغيير قيمة عامل التعلم اعتمادا على نسبة الخطأ الحالي للشبكه واستنتاج مقدار التغيير المطلوب على عامل التعلم للحصول على نتيجة افضل. ويوضح الشكل (4) تحسن اداء الشبكة من خلال وقت التدريب والسلوك العام للشبكه. ومن خلال دمج العاملين يمكن تحسين اداء الشبكه في دقة الحل الا انها قد تعطي فترة تدريب اكبر كما اظهرته نتائج البحث وموضح في الشكل (5).

References

[1] Chen, S., Chng, E.S., Alkadhimi, Regularized Orthogonal Least Squares Algorithm for Constructing Radial Basis Function Networks, Int J. Control, 1996, Vol. 64, No. 5, pp. 829-837.

[2] Benvenuto, N., & Piazza, F. (1992). On the complex backpropagation algorithm. IEEE Transactions on Signal Processing, 40(4), 967-969.

[3] Georgiou, G. M., & Koutsougeras, C. (1992), Complex domain back-propagation. IEEE Trans. Circuits and Systems--II: Analog and Digital Signal Processing, 39(5), 330-334.

[4] Scales, L. E., Introduction to Non-Linear Optimization, New York: Springer-Verlag, 1985.

[5] Hassoun, M., Fundamentals of Artificial Neural Networks, MIT, 1995.

[6] Kim, M. S., & Guest, C. C. (1990). Modification of backpropagation networks for complex-valued signal processing in frequency domain. Proceedings of the International Joint Conference on Neural Networks, 3, 27-31.

[7] Richard P. Lippmann, An introduction to computing with neural nets, ACM SIGARCH Computer Architecture News, v.16 n.1, p.7-25, March, 1988 [doi>10.1145/44571.44572]

[8] Nitta, T., & Furuya, T. (1991). A complex back-propagation learning. Transactions of Information Processing Society of Japan, 32(10), 1319-1329. in Japanese.

[9] Tohru Nitta, An extension of the back-propagation algorithm to complex numbers, Neural Networks, v.10 n.9, p.1391-1415, Nov. 1997