

Optimizing Multilayer Neural Networks by using Genetic Algorithm

استخدام الخوارزمية الجينية لتصميم وتدريب الشبكات العصبية متعددة الطبقات

Osman Ahmed Abdalla

Yahya Al-Murtadha,

Abdelrahman Osman Elfaki,

د. عثمان أحمد عبد الله محمد، د. يحي المرتضي، د. عبد الرحمن عثمان الفكي

University of Tabuk
Faculty of Computers and
Information Technology,
Information Technology
Department

University of Tabuk
Faculty of Computers and
Information Technology,
Computer Science
Department

University of Tabuk
Faculty of Computers and
Information Technology,
Computer Science
Department

جامعة تبوك كلية الحاسبات و تقنية المعلومات-قسم تقنية المعلومات-المملكة العربية السعودية-تبوك

osman_uofg@yahoo.com, y.murtadh@gmail.com,

abdelrahmanelfaki@gmail.com

المستخلص

تعتبر عملية تحديد واختيار المعلمات المثلى للشبكات العصبية وبخاصة للشبكات العصبية متعددة الطبقات عملية هامة و شاقة للغاية، وتكمن أهميتها بإعتبار النجاح في إختيار وتحديد هذه المعلمات يؤدي لتصميم شبكات عصبية ذات فاعلية ودقة في الأداء. هذا الورقة تستخدم الخوارزمية الجينية كطريقة مثلى لإختيار معلمات تصميم الشبكات العصبية وذلك بإستخدام الترميز الثنائي بهدف الحصول على تصميم شبكة عصبية فعالة وأقل اعتماداً على الإنسان. إستخدمت الخوارزمية الجينية لإختيار وتحديد المعلمات الاتية: عدد الطبقات المخفية للشبكة وعدد الخلايا في الطبقة او الطبقات المخفية، نوع خوارزمية التدريب، نوع دالة التفعيل، الوزن الاولي، معدل التعلم، معدل قوة الدفع، و عدد تكرارات التدريب. وأوضحت النتائج الأولية للطريقة المقترحة أن النموذج الجديد لتحديد معلمات تصميم وتدريب الشبكة العصبية أسهمت في تقديم أداء مرضي وفُعال وحلت مشكلة إهدار الوقت التي تظهر بإستخدام الطرق التقليدية.

1. المقدمة

تحديد وإختيار معلمات تصميم وتدريب الشبكات العصبية عملية هامة للغاية، والهدف المطوب من الاختيار الامثل هو تصميم شبكات عصبية أكثر دقة وأقل اعتماداً على الإنسان، تقديم أداء مرضي، والقدرة على اتمام عملية التدريب في وقت مناسب. الهدف من هذه الورقة هو اختيار معلمات تصميم وتدريب الشبكات العصبية وذلك بغرض تحسين كفاءة أداء الشبكة من خلال استخدام طريقة الترميز الثنائي للخوارزمية الجينية.

عموماً، يتأثر التطبيق الناجح للشبكات العصبية لغرض التنبؤ والنمذجة في مجالات العلوم والهندسة بشكل كبير على اختيار المعلمات التالية: بنية الشبكة (أي عدد الطبقات المخفية وعدد الخلايا العصبية في كل طبقة مخفية)، نوع خوارزمية التدريب (كمثال: التدرج

النسبي، فاس-نيوتن، المكورات التدريجية، ليفينبيرج-ماركوارت، الخوارزمية عكسية المرنة، الخ)، نوع دالة التفعيل (كمثال: اللوجستي السيني، الخطي، الخ)، معدل التعلم، معدل قوة الدفع، وعدد التكرارات لتدريب الشبكة.

الشبكات العصبية [3, 2, 1]، تستخدم بكثرة في النمذجة والتنبؤ وبخاصة لحل المسائل ذات النسق غير الخطي، التي تتغير مع الزمن، وصعوبة والتي يصعب إستنتاج أو تمثيل العلاقة بين مدخلاتها ومخرجاتها.

في الوقت الراهن تعتبر الخوارزمية الجينية واحدة من أهم طرق البحث عن التحسين اعتماداً على مبادئ الوراثة والتطور [6, 4, 5]. الأبحاث الحالية في الشبكات العصبية توضح أن تحديد المعلمات المثلى للتصميم والتدريب هي العقبة الرئيسية التي تواجه عملية تطوير الشبكات العصبية.

عدد الطبقات المخفية و عدد الخلايا العصبية في كل طبقة لها تأثير كبير في أداء تدريب الشبكة، مقدار وقت التدريب، والقدرة على التعميم. وعلاوة على ذلك، استخدام عدد كبير من الخلايا العصبية في الطبقة او الطبقات المخفية قد يتناسب من البيانات أكثر من المطلوب، مما قد يتسبب في فقدان القدرة على تعميم الشبكة. الى جانب ذلك، استخدام عدد قليل من الخلايا العصبية في طبقة او الطبقات المخفية يؤدي الي شبكة قد لا تكون قادرة على التعلم.

الهدف الرئيسي من هذا البحث هو تصميم شبكة عصبية مع أداء مرضي وأقل اعتماداً الإنسان، و قدرة على حل مشكلة الوقت المهدر في عملية التدريب.

في السنوات الأخيرة، تم تطبيق الشبكات العصبية بنجاح في العديد من الأعمال في مجالات العلوم و الهندسة، ولكن لا يوجد حتى الآن طريقة مثلى و عامة لتصميم الشبكات العصبية. تهتم هذه الورقة بشكل رئيسي على تصميم الشبكات العصبية من خلال إختيار المعلمات بصورة مؤتمتة بالكامل من خلال إستخدام الترميز الثنائي للخوارزمية الجينية. هنالك العديد من الأعمال التي تم تنفيذها في السنوات الأخيرة؛ والتي يمكن تلخيصها على النحو التالي:

أكثر استخدامات الخوارزمية الجينية لتحسين عمل الشبكات العصبية في المجال الهندسي كان في جانب اختيار شروط التشغيل الأمثل، إختيار واعداد معلمات المدخلات لتحقيق أكبر قدر من الأداء، و أيضاً زيادة و إنقاص المخرجات حسب الحاجة [7].

تم استخدام الخوارزمية الجينية في اختيار أوزان الاتصال الأولية والتحيز للشبكة العصبية ذات الانتشار الخلفي وذلك لتجنب التقارب في الحلول المثلى [9, 8].

لاختيار اوزان الاتصال الاولية و دالة التفعيل للشبكة العصبية استخدم الترميز الحقيقي للخوارزمية الجينية من اجل الحصول على اداء فعال كما تم مقارنة هذه الطريقة مع طرق اخري و اوضحت النتائج فعالية استخدام هجين الشبكات العصبية مع الخوارزمية الجينية [10].

تعتبر معمارية الشبكة من اهم معلمات تصميم الشبكات العصبية و تلعب دوراً كبيراً في تسريع اداء الشبكة لذلك تم استخدام الخوارزمية الجينية لاختيار امثل عدد للخلايا العصبية في الطبقة المخفية و اوضحت النتائج ان الشبكة المطورة كان ادائها افضل بكثير من غيرها [11].

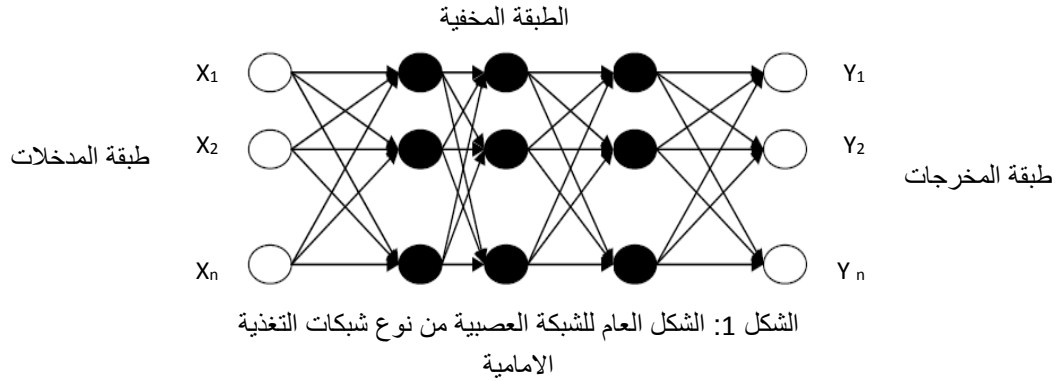
يلاحظ من مما قدم من بحوث مشابهة ركزت فقط على معالم محددة مما يستوجب تقديم بحث يغطي معظم معالم تصميم و تدريب الشبكات العصبية.

2. الشبكات العصبية

الغرض من الشبكات العصبية هو محاكاة العقل البشري في اتخاذ القرارات والتنبؤ ومحاكاته أيضاً من حيث البنية والوظيفة. في السنوات الأخيرة، تم تطبيق تقنية الشبكات العصبية في العديد من المجالات، مثل صناعة السيارات [7]، الخدمات المصرفية [8]، الإلكترونيات [9]، الخدمات المالية [10]، الصناعة [11]، الاتصالات [12]، النفط والغاز [13]، والروبوتات [14].

تتكون الشبكات العصبية من عدة طبقات، وهي: طبقة المدخلات، الطبقة المخفية، و طبقة المخرجات. هنالك العديد من انواع الشبكات العصبية لكن الشبكات العصبية ذات التغذية الأمامية هو الأكثر استخداماً وذلك لسهولةها و بساطة تكوينها، ودائماً ما تستخدم معها خوارزمية الانتشار العكسي كخوارزمية تدريب.

الشكل 1 التالي يوضح النموذج العام للشبكات العصبية ذات التغذية الأمامية.

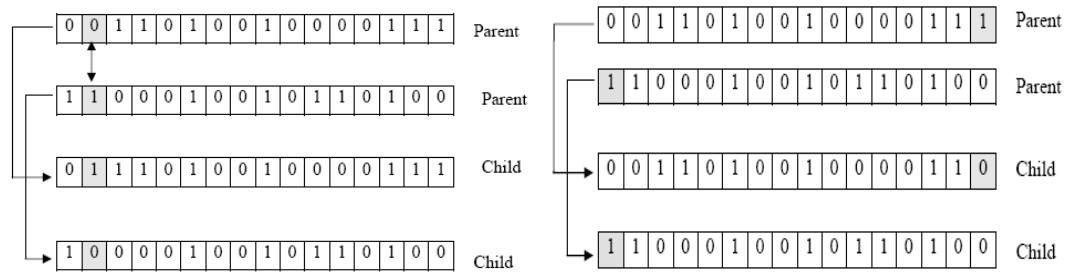


3. الخوارزمية الجينية

هو تقنية اندماجية مثلي وعامة الغرض تعتمد على نظرية داروين للنشوء و التطور و التي تبحث عن أقرب/أمثل قيمة لدالة الهدف عن طريق محاكاة عملية التطور الطبيعية. وقد أستخدمت الخوارزمية الجينية بنجاح لحل العديد من المشكلات في مجموعة واسعة من المجالات (غولدبرغ، 1989).

باختصار، الخوارزمية الجينية تقوم بثلاث عمليات أساسية وهي: الإختيار، التعابر، والطفرات. تبدأ الخوارزمية بمجموعة من الحلول للمشكلة التي يجري فحصها، مجموعة الحل (تمثل بالكروموسومات في الخوارزمية الجينية) تسمى بالسكان.

تستخدم عملية التعابر للحصول على حل جديد من خلال الجمع بين الكروموسومات المختلفة وذلك لتوليد ذرية او اجيال جديدة أفضل، تسمى عملية الحصول على حل جديد عن طريق تغيير عضو في السكان تسمى هذه العملية الطفرة. يمكن تمثيل هذه العمليات كما هو مبين في الشكلين 2 و 3.



الشكل 3: عملية الطفرة

الشكل 2: عملية التعابر

4. المنهجية

1-4 جمع البيانات

في هذا البحث استخدمنا نوعين من البيانات لتصميم وتدريب الشبكة، الاولى لبيانات جمعت من ركة بتروناس لتكرير البترول، الثانية لبيانات قياسية للمشكلة XOR.

تم جمع بيانات حجمها 500 بيانه من شركة بتروناس لتكرير البترول بمدينة ملاكا بماليزيا لتطوير وتدريب النموذج المقترح، تم فحص البيانات بصرياً واحصائياً للتأكد من صحتها وازالة القيم الشاذة وكذلك القيم الغير مكتملة. الشبكة المقترحة سوف تستخدم للتنبؤ بعزل البنتان isolate pentane (iC_5) و البنتان الطبيعي normal pentane (nC_5).

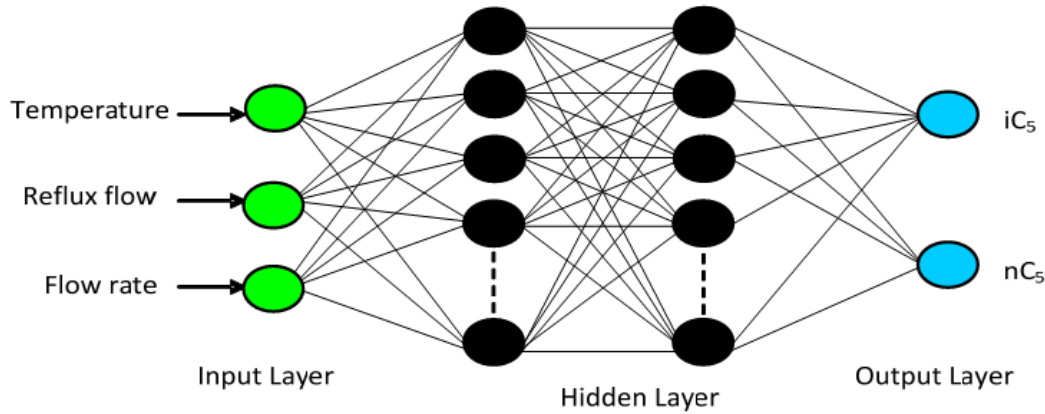
مدخلات هذه الشبكة هي المتغيرات التالية:

- درجة الحرارة Temperature
- انسياب البخار الراجع Reflux flow
- معدل الانسياب Flow rate

المخرجات هي:

- البنتان المعزول iC_5
- البنتان الطبيعي nC_5

تمثل هذه المتغيرات كطبقة ادخال و طبقة مخرجات كما هو موضح في الشكل 4 أدناه.



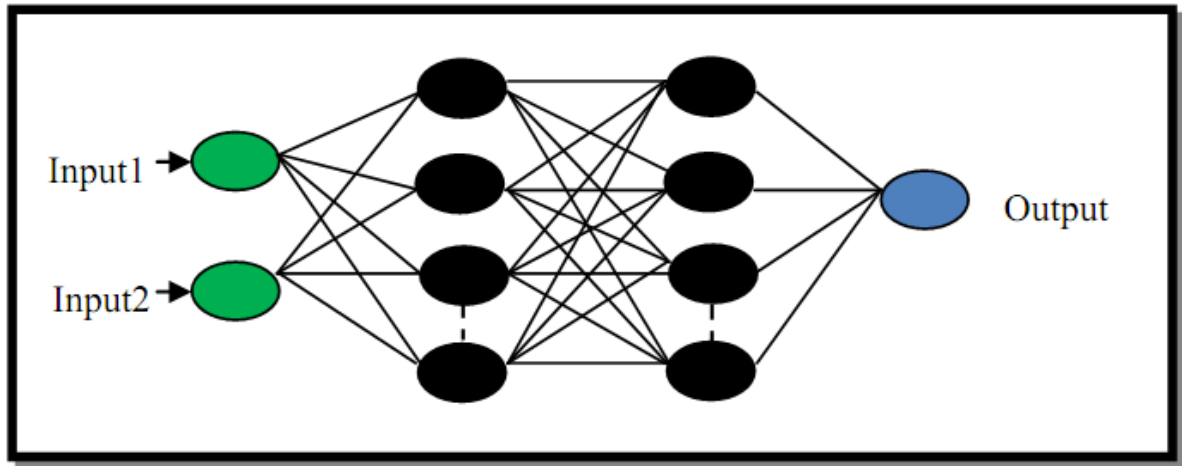
الشكل 4 نموذج الشبكة العصبية لمشكلة استخلاص البيوتان

عادة ما تستخدم XOR كمشكلة قياسية للتأكد من صحة العديد من الطرق الرياضية و المنطقية و غيرها؛ تحتوي المشكلة على متغيرين ثنائيين كمدخل و متغير ثنائي واحد كمخرج، واستخدمت سابقاً في العديد من الدراسات كمشكلة قياسية لاختبار أداء الشبكات العصبية. تتكون XOR من أربع عينات تدريب، اذا كانت المدخلات متطابقة فالناتج صفر، و اذا كانت المدخلات مختلفة فالناتج واحد. الجدول 1 يوضح جدول الحقيقة للمشكلة XOR.

الجدول 1 جدول الحقيقة للمشكلة XOR

Input1	Input2	Output= $x \oplus y$
1	1	0
1	0	1
0	1	1
0	0	0

الشكل 5 التالي يوضح بنية الشبكة العصبية متعددة الطبقات للمشكلة XOR.



الشكل 5 نموذج الشبكة العصبية للمشكلة XOR

2-4 مزيج الشبكة العصبية و الخوارزمية الجينية المقترح:

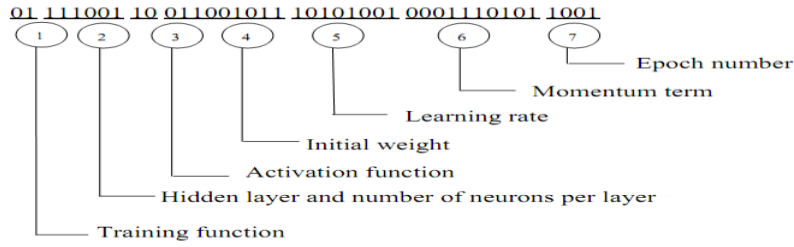
مزج الشبكة العصبية مع الخوارزمية الجينية يعتبر وسيلة فعالة لأغراض النمذجة والأمثلية. في هذا البحث، سوف يتم تطبيق الخوارزمية الجينية للحصول على معلمات تصميم وتدريب الشبكة العصبية المثلى.

1-2-4 ترميز الخوارزمية الجينية لمعلمات تصميم وتدريب الشبكة العصبية:

سيتم تمثيل معلمات التصميم والتدريب من خلال الترميز الثنائي، وذلك بتمثيل المعلمات في شكل كروموسومات، في حالتنا، سيتم تمثيل المعلمات التالية:

- نوع خوارزمية التدريب يمثل بثنائيتين.
- بنية الشبكة (عدد الطبقات المخفية و عدد الخلايا في كل طبقة مخفية) تمثل بست ثنائيات.
- نوع دالة التفعيل في خلايا الطبقة المخفية و طبقة المخرجات يمثل بثنائيتين.
- الوزن الاولي يمثل بعشر ثنائيات.
- معدل التعلم يمثل بثمانية ثنائيات.
- معدل قوة الدفع يمثل بعشر ثنائيات.
- حجم تكرار التدريب يمثل باربع ثنائيات.

النمط الوراثي هو سلسلة من الثنائيات (0 أو 1) بطول ثابت. كل نمط وراثي يتوافق مع كروموسوم فريد. الشكل 6 يوضح طريقة تمثيل معلمات التصميم والتدريب للشبكة العصبية من خلال الترميز الثنائي.



الشكل 6: تمثيل معلمات التصميم والتدريب بواسطة الترميز الثنائي

في هذا الجزء سنتعرف على نمطين حواريين للتدريب هما العملية السميكية بواسطة الترميز الساسي:

تعتبر تقنية الانتشار الخلفي من أكثر خوارزميات تدريب الشبكات العصبية استخداماً و تستخدم لتعديل أو ضبط اوزان اتصال الشبكة لتقليل نسبة الخطأ خلال عملية التدريب [22]. في هذا البحث استخدمنا اربع من أشهر انواع خوارزميات تدريب الشبكة وهي: قاوس-نيوتن، المكورات التدرجية، ليفينبيرج-ماركوارت، الخوارزمية عكسية المرونة.

الجدول 2 يوضح تمثيل هذه الخوارزميات بواسطة الترميز الثنائي.

جدول 2 تمثيل خوارزميات التدريب

تسلسل الثنائية	خوارزمية التدريب
00	المكورات التدرجية
01	قاوس- نيوتن
10	ليفينبيرج-ماركوارت
11	الخوارزمية عكسية المرونة

وكذلك مُثلت معمارية الشبكة وتعني عدد الطبقات المخفية و الخلايا التي تحتويها اي طبقة بعدد ست ثنائيات كمثال: التسلسل الثنائي 000000 يعني طبقة مخفية واحدة تحتوي علي 3 خلايا. وعلى نفس المنوال مُثلت نوع دالة التفعيل، معدل التعلم، معدل قوة الدفع و حجم تكرار التدريب.

3-4 تقييم الملائمة

في هذا البحث استخدمنا الصيغة الرياضية أدناه لتقييم الملائمة والتي ايضاً استخدمت بواسطة [33]:

$$\text{Fitness} = 10^6 - \frac{1}{n} \sum_{i=1}^n (y_i - y_t)^2 \quad (1)$$

حيث y_t تشير للمخرج المرغوب فيه و y_i تشير للمخرج الخالي.

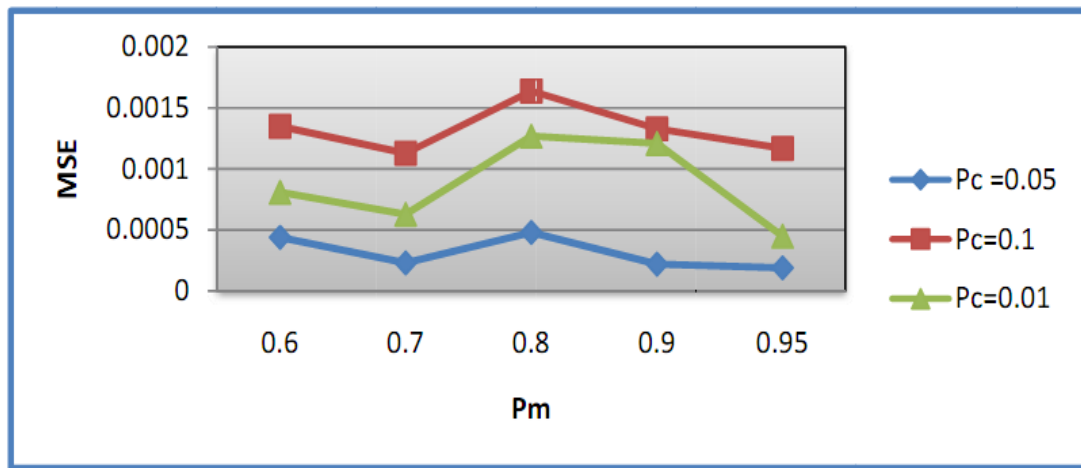
4-4 مقاييس الأداء

متوسط مربع الخطأ استخدم لقياس كفاءة أداء الشبكة وفق الصيغ التالية:

$$MSE = \frac{1}{n} \sum_{i=1}^k (y_i - \hat{y}_i)^2 \quad (2)$$

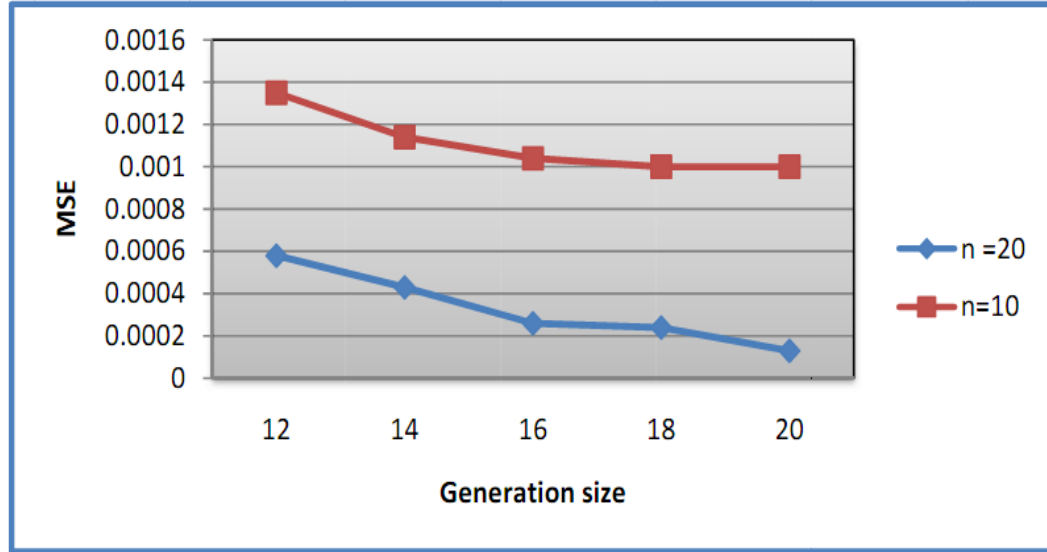
5. النتائج:

الشكل 7 يوضح أفضل متوسط مربع الخطأ الذي حققه النموذج المقترح من خلال التدريب باختيار ثلاث قيم مختلفة لاحتمال انتقال الطفرة. وقد تحقق أفضل أداء لقيمة احتمال التعابر = 0.95 وقيمة احتمال الطفرة = 0.05.



الشكل 7 متوسط مربع الخطأ على حسب احتمال الطفرة و احتمال التعابر

في هذه الدراسة تم اختبار عدد من الأفراد ما بين 10 و 20 من عدد السكان، يوضح الشكل 5 أفضل لقيم متوسط مربع خطأ، أفضل أداء تم الحصول عليه عند حجم السكان 20، تم اختبار قيم مختلفة من الأجيال كمثل 10 و 20 و تحقق أفضل أداء بعد الجيل الـ 20.



الشكل 8 متوسط مربع الخطأ على حسب حجم الاجيال و السكان

كمثال للتناج لمشكلة استخلاص البيوتان حصلنا على افضل حل وكان التسلسل كالاتي:

10 001001 01 100000001 00000100 0000111110 0001

ويفسر التسلسل الثاني كالاتي:

الشبكة عصبية المثلي لها المعلمات الاتية:

- طبقة مخفية واحدة تحتوي على 15 خلية عصبية
- نوع دالة التنفيع اللوجستية السينية للخلايا العصبية المخفية و الدالة الخطية للخلايا العصبية للمخرجات
- خوارزمية التدريب المكورات التدريجية
- وزن الاتصال الاولي = 0.519716
- معدل التعلم = 0.025857471
- معدل قوة الدفع = 0.061401
- بتكرار تدريب قدره 200

هذا الحل حقق متوسط مربع خطأ قدره 0.0019 لمخرج IC₅ و 0.0002 للمخرج nC₅

كمثال للتناج للمشكلة XOR حصلنا على افضل حل وكان التسلسل كالاتي:

00 111100 10 0000011010 00000110 0000110001 0011

ويفسر التسلسل الثاني كالآتي:

الشبكة عصبية المثلي لها المعلمات الآتية:

- طبقتان مخفيتان في الاولي 8 خلايا عصبية وفي الثانية 5 خلايا عصبية.
- نوع دالة التفعيل السينية الظل للخلايا العصبية المخفية و للخلايا العصبية للمخرجات
- خوارزمية التدريب ليفينبيرج-ماركوارت
- وزن الاتصال الاولي = 0.059031
- معدل التعلم = 0.031922630
- معدل قوة الدفع = 0.052192
- بتكرار تدريب قدره 400
- هذا الحل حقق متوسط مربع خطأ قدره 0.0002 .

6. الخاتمة:

في هذه الدراسة قدمنا نموذج يستخدم الترميز الثنائي للخوارزمية الجينية لإختيار امثل معلمات تصميم وتدريب الشبكات العصبية. طبقنا النموذج المقترح علي مشكلة حقيقية وهي مشكلة استخلاص البيوتان، واخري قياسية وهي المشكلة القياسية XOR وحققت الطريقة المقترحة نتائج مرضية وفعالة. تتميز الدراسة عن غيرها بانها ركزت علي معظم معلمات التصميم والتدريب كاختيار بنية الشبكة، نوع خوارزمية التدريب، نوع دالة التفعيل، الوزن الاولي، معدل التعلم، معدل قوة الدفع و حجم تكرار التدريب.

المراجع:

- [1] H. Patrick C.L. Application of artificial neural networks to the prediction of sewing performance of fabrics”, International Journal of clothing and Technology, 2007, Vol. 19, No.5, pp. 291318.
- [2] F. Fred F., G. James D., and L. Juliet N., Predicting temperature profiles in producing oil wells using artificial neural networks, Engineering Computation, 2000, Vol. 17, No. 6, pp. 0264-4401.
- [3] K. Okyay, Springer, Artificial Neural Networks and Neural Information, 2003, ISBN 3540404082.
- [4] D. Satyanarayana, K. Kamarajan, and M.Rajappan (). “Genetic Algorithm Optimized Neural Networks Ensemble for Estimation of Mefenamic Acid and Paracetamol in Tablets”, Genetic Algorithm Optimized Neural Networks Ensemble, Acta Chim. Slov., 2005, Vol. 52, pp. 440–449.
- [5] M. Izadifar, M. Zolghadri Jahromi, Application of genetic algorithm for optimization of vegetable oil hydrogenation process, Journal of Food Engineering, 2007, Vol. 78, Issue 1, pp. 1-8.
- [6] F. Konstantinos P., Biological engineering applications of feedforward neural networks designed and parameterized by genetic algorithms, Neural Networks, 2005, Vol. 18, Issue 7, pp. 934-950.
- [7] I. Istadi, N.A.S. Amin, Chemical Engineering Science, 2007, Vol. 62, pp. 6568
- [8] Ming Tan, Gaohong He, Fei Nie, Lingling Zhang, Liangping Hu, Optimization of ultrafiltration membrane fabrication using backpropagation neural network and genetic algorithm, Journal of the Taiwan Institute of Chemical Engineers, 2014, Vol. 45, pp. 68–75
- [9] Fang Feng Ping, Fang Xue Fei, Multivariant Forecasting Mode of Guangdong Province Port throughput with Genetic Algorithms and Back Propagation Neural Network, Procedia - Social and Behavioral Sciences 2013, Vol. 96, pp. 1165 – 1174.
- [10] Rahman Ashena, Jamshid Moghadasi, Bottom hole pressure estimation using evolved neural networks by real coded ant colony optimization and genetic algorithm, Journal of Petroleum Science and Engineering, 2011, Vol. 77, pp. 375–385.

- [11] Meysam Mirarab, Maysam Sharifi, Mohammad Ali Ghayyem, Fatemeh Mirarab, Prediction of solubility of CO₂ in ethanol-[EMIM] [Tf₂N] ionic liquid mixtures using artificial neural networks based on genetic algorithm, *Fluid Phase Equilibria*, 2014.
- [12] L. Bor-Ren and R.G. Hoft, Neural networks and fuzzy logic in power electronics, *Control Engineering Practice*, 2003, Vol. 2, Issue 1, pp. 113-121.
- [13] Z. Xiaotian, X. Hong Wang, Li, and L. Huaizu Li, Predicting stock index increments by neural networks: The role of trading volume under different horizons, *Expert Systems with Applications*, 2008, Vol. 34, Issue 4, pp. 3043-3054.
- [14] G.R. Cheginia, J. Khazaeia, B. Ghobadianb, and A.M. Goudarzi, Constructing ANN, Prediction of process and product parameters in an orange juice spray dryer using artificial neural networks *Journal of Food Engineering*, 2008, Vol. 84, Issue 4, pp 534-543.
- [15] N. Perambur S. and A. Preechayasomboon, Development of a neuroinference engine for ADSL modem applications in telecommunications using an ANN with fast computational ability, *Neurocomputing*, 2002, Vol. 48, Issues 1-4, pp. 423-441.
- [16] F. Fred F., G. James D., and L. Juliet N. , Predicting temperature profiles in producing oil wells using artificial neural networks, *Engineering Computation*, 2000, Vol. 17, No. 6, pp. 0264-4401.
- [17] N. Huang, K.K. Tan, and T.H. Lee, Adaptive neural network algorithm for control design of rigid-link electrically driven robots, *Neurocomputing*, 2008, Vol. 71, Issues 4-6, pp 885-894.