SEBHA UNIVERSITY JOURNAL OF PURE & APPLIED SCIENCES VOL.22 NO. 3 2023
DOI: 10.51984/JOPAS.V22I3.2756

مجلة جامعة سبها للعلوم البحتة والتطبيقية
Sebha University Journal of Pure & Applied Sciences

Journal homepage: www.sebhau.edu.ly/journal/index.php/jopas

استخدام الخوارزميات الجينية لضبط وتحسين المعاملات الفائقة لشبكات العصبية

بدر نجيب عويدات

كلية تقنية المعلومات، جامعة الزيتونة، ترهونة، ليبيا

الملخص

توصلت هذه الدراسة إلى أن الخوارزمية الجينية تلعب دورًا حاسمًا في ضبط وتحسين المعاملات الفائقة لشبكات العصبية.MLP تم استخدام المعاملات الرئيسية، وهي حجم الطبقات الخفية، وعدد الدورات (الحقب)، ومعدل التعلم، وخوارزمية التحسين، لتحقيق أداء محسن للشبكة العصبية. تم استخدام مجموعة بيانات زهرة اللوتس ومجموعة بيانات الأرقام المكتوبة بخط اليد لتقييم أداء النموذج. وقد تم الحصول على نسبة دقة تبلغ 100% لمجموعة بيانات زهرة اللوتس، مما يشير إلى تحقيق أداء ممتاز للتصنيف. وبالنسبة لمجموعة بيانات الأرقام المكتوبة بخط اليد، تم الحصول على نسبة دقة تبلغ 94%، مما يدل على أداء جيد وفعالية النموذج في التعرف على الأرقام. باستخدام الخوارزمية الجينية، تم تحسين المعاملات الفائقة للشبكة العصبيةMLP، مما أدى إلى تحسين الأداء وزيادة دقة التصنيف في مجموعتي البيانات المستخدمتين. تؤكد هذه النتائج أهمية وفاعلية الخوارزمية الجينية في تحسين ضبط المعاملات الفائقة للشبكات العصبية، مما يمكن من تحسين أداء النماذج وتحسين قدرتها على التعرف على الأنماط والتصنيف بدقة.

الكلمات المفتاحية:

دقة التصنيف
شبكة عصبية MLP
مجموعة بيانات الأرقام المكتوبة بخط اليد
مجموعة بيانات زهرة اللوتس،
تحسين المعاملات الفائقة
خوارزمية جينية

Utilizing Genetic Algorithms for Tuning and Enhancing Hyperparameters of Neural Networks

Bader N. Awedat

Faculty of Information Technology, Al-Zay Zaytouna University, Tarhuna, Libya

Keywords:

Genetic algorithm
Hyperparameter optimization
MLP neural network
Classification accuracy
Iris dataset
handwritten Numeral dataset

ABSTRACT

The present study utilized a genetic algorithm to optimize the hyperparameters of an MLP neural network. The key hyperparameters, including the hidden layer sizes, number of epochs, learning rate, and optimization algorithm, were subjected to optimization using the genetic algorithm. The Iris flower dataset and the handwritten Numeral dataset were employed to evaluate the performance of the optimized model. The results showed significant improvements in the classification accuracy achieved through the genetic algorithm optimization. The Iris dataset yielded a perfect accuracy of 100%, indicating excellent classification performance. Similarly, the handwritten Numeral dataset achieved an accuracy of 94%, demonstrating the effectiveness of the optimized model in recognizing handwritten Numeral. These findings underscore the importance of genetic algorithms in fine-tuning and enhancing the hyperparameters of MLP neural networks. The application of the genetic algorithm resulted in improved performance and increased classification accuracy in both datasets, showcasing the potential of this approach in improving the performance of neural network models.

المقدمة

على الرغم من استخدام الشبكات العصبية الاصطناعية لحل مجموعة متنوعة من المشكلات، إلا أنها لا تزال تعاني من بعض القيود. واحدة من أكثر القيود شيوعًا ترتبط بتدريب الشبكات العصبية. تعتبر خوارزمية التعلم بالانتشار العكسي (Back-Propagation) شائعة الاستخدام بفضل مرونتها وقابليتها للتطبيق الرياضي. ولكن من عيوبها انه لا يمكن ضمان الوصول للحل الأمثل في التطبيقات العملية، وقد تتقارب خوارزمية الانتشار العكسي نحو مجموعة من الأوزان بشكل غير مثالي. نتيجة لذلك، غالبًا ما تكون هذه الخوارزمية غير قادرة على إيجاد حل جيد للمشكلة المطروحة. وتوجد صعوبة أخرى تتعلق باختيار هيكلية او طوبولوجيا مثلى للشبكة العصبية لان بناء الهيكل المناسب لشبكة العصبية في مشكلة معينة يتم غالبًا بالطرق

*Corresponding author:
E-mail addresses: bader_najep@yahoo.com
Article History : Received 17 July 2023 - Received in revised form 21 September 2023 - Accepted 02 October 2023

التقليدية، ولا يزال تصميم طوبولوجيا الشبكة العصبية فنًا أكثر منه هندسة.

الخوارزميات الجينية هي تقنية تحسين فعالة يمكن أن توجه الي تحسين الاوزان واختيار طوبولوجيا الشبكة العصبية بشكل فعال واكثر مثالية. تساعد الخوارزميات الجينية في اختيار طوبولوجيا شبكة معينة. حيث ان الفكرة الأساسية وراء تطور هيكلية الشبكة المناسبة هي إجراء بحث جيني في مجموعة من الهياكل الممكنة. يجب أن نختار أولاً طريقة لترميز هندسة الشبكة في صبغة. هناك العديد من الطرق المختلفة لترميز هيكلية الشبكة. المفتاح هو تحديد كمية المعلومات المطلوبة لتمثيل الشبكة. كلما زادت معلمات هندسة الشبكة، زادت تكلفة الحساب. في هذا البحث سوف نقوم باستخدام الخوارزميات الجينية للوصول الي طوبولوجيا مثالية للشبكة العصبية، حيث تكمن أهمية البحث في التركيز على اختيار الطبقات القابلة للتدريب (trainable layers) في هذه الشبكات والتي تعتبر من المهام الرئيسية لتصميم النماذج. والتي عادة ما يتم تصميم نماذج الشبكات العصبية التقليدية يدويًا بناءً على الحدس والخبرة.

الدراسات السابقة

1. تناولت دراسة تشين لي واخرون (2022) مشكلة تحسين المعاملات الفائقة (hyperparameters) في تطوير شبكات عصبية عميقة بتقنية نقل التعلم (Transfer Convolutional Neural Network). تم استخدام خوارزمية جينية (Genetic Algorithm) لاختيار الطبقات القابلة للتدريب في النموذج النقال للتعلم. تم التركيز على معيار الدقة وعدد الطبقات القابلة للتدريب. أظهرت النتائج أن الطريقة كفؤة في هذه المهمة، حيث تم تحقيق نسبة بدقة تصل إلى 97% في تصنيف مجموعات البيانات للقطط والكلاب، وذلك في ما لا يزيد عن 15 جيل [1].

2. قالت دراسة جي هون هان واخرون (2020) ان مشكلة تحسين المعلمات الفائقة في تطوير خوارزميات التعلم العميق صعبة للغاية. تم استخدام خوارزمية جينية محسنة لحل هذه المشكلة. تم اعتبار الدقة ووقت التحقق كمعايير لتقييم الأداء. تم تقييم الخوارزمية باستخدام نموذجين يتألف النموذج الأول من طبقة تكاملية واحدة وطبقة واحدة متصلة بالكامل، والنموذج الثاني متعدد الطبقات مكون من ثلاثة طبقات الطبقة الأولى طبقة تكاملية والطبقة الثانية طبقة متصلة بالكامل والطبقة الثالثة طبقة مخرجات . تم استخدام مجموعة بيانات MNIST ومجموعة بيانات تشخيص أعطال المحرك لتدريب الخوارزمية. أظهرت النتائج أن الشبكة العصبية ذات الطبقة الواحدة قادرة على تحقيق معدل دقة 99.2% على مجموعة بيانات MNIST، في حين أظهرت النتائج أن الشبكة العصبية متعددة الطبقات قادرة على تحقيق معدل دقة 98.5% على مجموعة بيانات تشخيص أعطال المحرك وقد أعطت الشبكة العصبية ذات الطبقة الواحدة انخفاضًا في وقت التدريب بنسبة 20%. بينما أظهرت الشبكة العصبية متعددة الطبقات انخفاضًا في وقت التدريب بنسبة 30% [2] .

3. توضح دراسة سينان إردن (2023) أهمية توقع التلوث الهوائي بدقة نظرًا لتأثيره السلبي على صحة الإنسان وتسببه في أمراض خطيرة، وذلك في سياق الاستدامة البيئية. على الرغم من استخدام الأساليب الإحصائية وتعلم الآلة التقليدية على نطاق واسع في توقع جودة الهواء، إلا أنها تعاني من قيود في اكتشاف العلاقات غير الخطية ونمذجة البيانات التسلسلية. في السنوات الأخيرة، تم تطبيق طرق التعلم العميق بنجاح في العديد من المجالات البحثية. في هذه الدراسة، تم استخدام خوارزمية التعلم العميق لتوقع مجموعة بيانات PM2.5، والتي تتضمن ملوثات الهواء (NO ، NO2، NOX، O3، PM2.5، SO, SO2) والميزات المناخية (سرعة الرياح، اتجاه الرياح، ودرجة حرارة الهواء) في منطقة اسطنبول الكبرى. تحتوي خوارزميات التعلم العميق على العديد من المعلمات الهامة مثل معدل التعلم ومعدل الإسقاط وعدد الطبقات الخفية ووحدات كل طبقة خفية ووظيفة التنشيط ووظيفة الخسارة وأداة التحسين التي يجب تحسينها من أجل تحقيق أداء التدريب الأمثل. لذلك، يتم اقتراح نهج تحسين المعلمات الهجينة بناءً على خوارزمية جينية لإيجاد أفضل تركيبة للمعلمات. تمت مقارنة نتائج توقع خوارزميات التعلم العميق مع المعلمات الافتراضية وخوارزميات البحث العشوائي للتأكد من فعالية نهج خوارزمية الجينات. أظهرت الطريقة المقترحة أداءً أفضل من التكوينات الأخرى، حيث تم تقليل خطأ MSE بنسبة 13.38% للمعلمات الهجينة واعطت المعلمات الافتراضية نسبة 55.30% لمقياس MSE. كشفت النتائج التجريبية أن خوارزميات الجينية واعدة وقابلة للتطبيق في تحسين المعلمات الفائقة لنماذج الشبكات العصبية العميقة، وخاصة في توقع جودة الهواء [3].

4. قال سانغهيوب لي واخرون (2021) ان الزهايمر هو أحد التحديات الرئيسية لشيخوخة، وتشخيصه وتوقعه من خلال مؤشرات حيوية مختلفة هو الأمر الرئيسي. في حين انتشرت تقنيات التعلم العميق كتقنيات في صناعة الرعاية الصحية، إلا أن تصميمها التجريبي صعب جدًا. والسبب الرئيسي لهذه المشكلة هو أن أداء الشبكات العصبية التلافيفية (CNN) يختلف بشكل كبير اعتمادًا على التوزيع الإحصائي لمجموعة بيانات الدخل. وتؤثر المعلمات الفرعية المختلفة بشكل كبير أيضًا على تقارب نماذج الشبكة العصبية التلافيفية. تم تحسين هيكل الشبكة باستخدام الخوارزمية الجينية استخدمت مجموعة بيانات صور المخ الأميلويدية التي تستخدم لتشخيص مرض الزهايمر. واعطت الخوارزمية تفوق بنسبة 11.73% في مهمة التصنيف المحددة على الخوارزميات التقليدية التي تم مقارنتها بها. كما تبين أن الشبكة المحسنة أكثر ثباتًا من الشبكات العصبية الأخرى، حيث كانت أقل عرضة للتأثر بالتغيرات في التوزيع الإحصائي لبيانات الإدخال [4].

الجدول 1: ملخص الدراسات السابقة

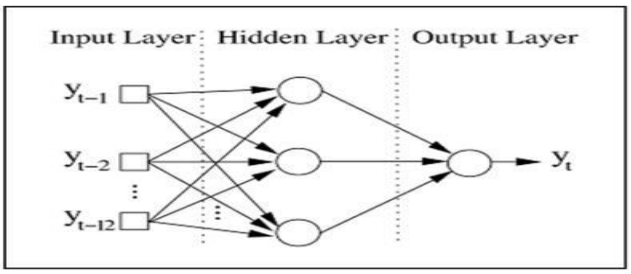
الدراسة	المشكلة	المنهجية	المعايير	النتائج الرئيسية
تشين لي وآخرون (2022)	تحسين المعاملات الفائقة في شبكات التعلم العميق بتقنية نقل التعلم	استخدام خوارزمية جينية لاختيار الطبقات القابلة للتدريب	الدقة وعدد الطبقات القابلة للتدريب	نسبة دقة تصل إلى 97% في تصنيف مجموعات بيانات القطط والكلاب
جي هون هان وآخرون (2020)	تحسين المعاملات الفائقة في خوارزميات التعلم العميق	استخدام خوارزمية جينية لتحسين الدقة ووقت التحقق	الدقة ووقت التحقق	معدل دقة 99.2% على مجموعة بيانات MNIST لنموذج ذو الطبقة الواحدة، ومعدل دقة 98.5% على مجموعة بيانات تشخيص أعطال المحرك لنموذج متعدد الطبقات
سينان إردن (2023)	توقع جودة الهواء باستخدام التعلم العميق	استخدام خوارزمية جينية لتحسين معلمات الشبكة العصبية	الدقة ووقت التحقق	تقليل خطأ MSE بنسبة 13.38% باستخدام المعلمات الهجينة
سانغهيوب لي واخرون (2021)	تحسين معاملات شبكات التعلم العميق لتشخيص الزهايمر	استخدام خوارزمية جينية لتحسين هيكل الشبكة	الدقة	تفوق بنسبة 11.73% في مهمة التصنيف المحددة للمجموعة البيانات المستخدمة

المواد وطرق العمل

في هذه الدراسة، تم بناء الخوارزميات باستخدام لغة برمجة Python وإطار عمل التعلم العميق Keras على منصة Google Colab. والتي تعتبر أداة محاكاة لإجراء التجارب وإنشاء النماذج، وهو متاح ومجاني للباحثين. تم استخدام مكتبة TensorFlow في برمجة وتنفيذ خوارزمية MLP وتدريبها. استخدمت مجموعتي بيانات الأولى هي مجموعة بيانات MNIST والتي تحتوي على 70,000 صورة رقمية لأرقام مكتوبة يدويًا بحجم 28x28 بكسل. تم تقسيم مجموعة البيانات إلى 60,000 صورة للتدريب و 10,000 صورة للاختبار، والثانية مجموعة بيانات زهرة اللوتس والتي تحتوي على 150 سجل لثلاثة أنواع من زهرة اللوتس وتم تقسيمها إلى مجموعة تدريب ومجموعة اختبار بنسبة 80:20 على التوالي. تم تنفيذ الخوارزمية الجينية على مستوى المعاملات الفائقة (hyperparameters) لخوارزمية MLP . تم تعريف دالة التطور الجيني لتكوين مجتمع من الخوارزميات المتعددة الطبقات باستخدام المكونات المختلفة للتطور الجيني، مثل حجم الطبقات الخفية وعدد الدورات (الحقب) و معدل التعلم وخوارزمية التحسين.

الشبكات متعدد الطبقات (Multilayer Perceptron):

هي نوع من أنواع الشبكات العصبية الاصطناعية التي تلقت اهتمام كبير من طرف الباحثين، وهذا يرجع لبساطتها وتطبيقاتها الناجحة في مختلف المجالات، تتكون شبكة MLP البسيطة من طبقة الإدخال وطبقة اخراج واحدة وتتضمن طبقة مخفية على الأقل، تقوم عملية التعلم لشبكة MLP بشكل عام عن طريق خوارزمية الانتشار الخلفي التي تعتبر الأكثر استخداما[9] .



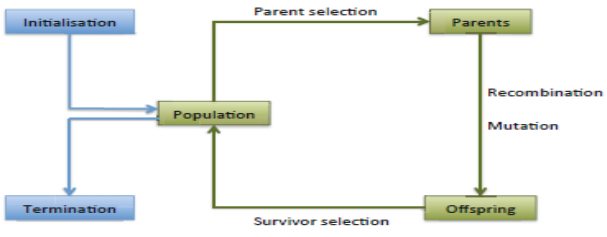
الشكل 1: شبكة MPL [9].

الخوارزميات الجينية:

تعتمد الخوارزميات الجينية على النهج التطوري لتعلم الآلة والذي يعتمد على النماذج الحسابية للانتقاء الطبيعي والوراثة. ويطلق عليه الحساب التطوري. وهو مصطلح شمولي يجمع بين الخوارزميات الجينية واستراتيجيات التطور والبرمجة الجينية. تحاكي كل هذه التقنيات التطور باستخدام عمليات الانتقاء (Selection) والتحور (Mutation) والتكاثر (Reproduction).

بشكل عام في الخوارزميات الجينية، يتم استخدام عملية التزاوج (Recombination/Crossover) للجمع بين صفات الوالدين وإنتاج نسل جديد يحمل مزيجًا من هذه الصفات. تتم عملية التزاوج عن طريق تقسيم الوالدين إلى جزئين وتبادل الأجزاء بينهما. عادةً ما يتم تنفيذ عملية التزاوج على نسبة محددة من الأفضل لياقة في السكان. تهدف عملية التزاوج إلى توليد تنوع واكتشاف حلول محتملة أفضل. يمكن تنفيذ عملية التزاوج باستخدام أساليب مختلفة مثل Single-Point Crossover و Multi-Point Crossover و Uniform Crossover عملية التزاوج تساهم في تحسين قدرة الخوارزمية الجينية على استكشاف وتحسين الحلول الممكنة.

بعد عملية التزاوج، يتعين اختيار الأفراد الناجين (Survivor Selection) لتشكيل الأجيال القادمة. تختلف استراتيجيات اختيار الناجين وفقًا لأهداف الخوارزمية، ولكن الاختيار الطبيعي-(Fitness Based Selection) هو الاستراتيجية الشائعة. يتم اختيار الأفراد ذوي اللياقة العالية للبقاء في السكان الجدد، مما يعزز التكيف وتحسين الحلول. تشمل استراتيجيات اختيار الناجين أيضًا عناصر عشوائية لتعزيز التنوع وتجنب التقطع الجيني كما هو موضح في الشكل (2) [5].



الشكل 2: المخطط العام للخوارزمية التطورية [5].

باختصار، في الخوارزميات الجينية عملية التزاوج تستخدم لجمع صفات الوالدين وإنتاج نسل جديد، بينما عملية اختيار الناجين تقوم بتحديد الأفراد الأكثر تكيفًا للبقاء في الأجيال القادمة. العمليتين تساهمان في تحسين الحلول وتكييف السكان مع البيئة المحيطة. في الشكل (3) تمثيل مبسط لخطوات الخوارزمية الجينية باستخدام الكود الزائف.

```
01- BEGIN
02-     INITIALIZE a population of candidate solutions;
03-     EVALUATE each candidate;
04-     REPEAT UNITL stop condition;
05-         SELECT parents;
06-         CROSS pairs of parents to generate new individuals;
07-         MUTATE new individuals;
08-         EVALUATE new population;
09-     END
10- END
```

الشكل 3: المخطط العام للخوارزمية التطورية بالكود الزائف [5].

هناك العديد من أنواع خوارزميات التطور الجيني، وتختلف في الطريقة التي تتعامل بها مع عمليات التطور والتكيف. بعض الأنواع الشائعة لخوارزميات التطور الجيني:

1. الخوارزمية الجينية الأساسية (Basic Genetic Algorithm):
تعتبر الخوارزمية الجينية الأساسية النموذج الأساسي لخوارزميات التطور الجيني. تستخدم عمليات التبادل العشوائي والتحور لتوليد تنوع جديد في الأجيال اللاحقة، وتستخدم وظيفة التكيف لتقييم فعالية الأفراد في البيئة.

2. الخوارزمية الجينية المعبأة (Packed Genetic Algorithm) :
تستخدم الخوارزمية الجينية المعبأة للتعامل مع المشكلات التي تتطلب تمثيل حلول معقدة وهياكل بيانات غير قياسية. يتم تعبئة الحلول المعقدة داخل الجينات باستخدام تقنيات مثل الترميز المرتب أو الترميز العشوائي.

3. الخوارزمية الجينية التوافقية (Fitness-based Genetic Algorithm):
تستخدم الخوارزمية الجينية التوافقية وظيفة التكيف المحددة بوضوح لتحقيق تطور الحلول نحو الأفضل. تتم استخدام أساليب مثل التحسين المتقدم والمحافظة على التنوع لتعزيز عملية التحسين والتكيف.

4. الخوارزمية الجينية متعددة الهدف (Multi-Objective Genetic Algorithm):
تستخدم الخوارزمية الجينية متعددة الهدف للمشكلات التي تتطلب تحقيق أكثر من هدف في الحل النهائي. تستخدم تقنيات مثل السياق البيئي والسياق الاقتصادي لتحقيق التوازن بين الأهداف المتعددة وتوليد مجموعة من الحلول الغير متفقة.

في هذا البحث تم استخدام الخوارزمية الجينية الأساسية والتي تستند على الخطوات التالية:

الخطوة 1: تمثيل الكروموسوم حيث يتم تمثيل الحلول المحتملة للمشكلة في شكل كروموسومات. في اغلب الاحيان يتم استخدام سلاسل ثنائية (بت-بايت) لتمثيل الجينات في الكروموسومات. يتم تعيين كل جين لقيمة محددة تمثل خاصية أو خيار من الحلول المحتملة.

الخطوة 2: دالة اللياقة (Fitness Function) وتستخدم لقياس جودة الحلول الممثلة بواسطة الكروموسومات. تحسب اللياقة بناءً على أداء الحلول الممثلة على مجموعة بيانات الاختبار أو بناءً على المعايير المحددة للمشكلة. الهدف هو زيادة اللياقة للوصول إلى أفضل الحلول.

الخطوة 3: الاختيار (Selection) حيث يتم اختيار الكروموسومات ذات اللياقة العالية للمشاركة في التكاثر. يتم تطبيق أساليب مختلفة للاختيار مثل الاختيار العشوائي، اختيار العينة المتساوية، اختيار العينة الأفضل، إلخ.

الخطوة 4: التكاثر (Reproduction) ويتم إنشاء نسل جديد من الكروموسومات المختارة في مرحلة الاختيار. يتم ذلك عن طريق تنفيذ عمليات المشغل الجيني مثل التبادل العشوائي (Crossover) والتحور (Mutation) لإنشاء تنوع وتطور جيل جديد من الكروموسومات.

يعتبر التبادل العشوائي (Crossover) المسؤول عن تبادل أجزاء من اثنين من الكروموسومات الفردية لإنشاء كروموسومات نسل جديدة. حيث يتم اختيار نقطة تبادل عشوائية في كل كروموسوم. ثم يتم استبدال الجزء بعد نقطة التبادل في الكروموسوم الأول بالجزء بعد نقطة التبادل في الكروموسوم الثاني، والعكس بالعكس. وفق الصيغة الرياضية[7]:

$$Offspring1 = Parent1[0: ExchangePoint] + Parent2[ExchangePoint + 1:] \quad (1)$$

$$Offspring2 = Parent2[0: ExchangePoint] + Parent1[ExchangePoint + 1:] \quad (2)$$

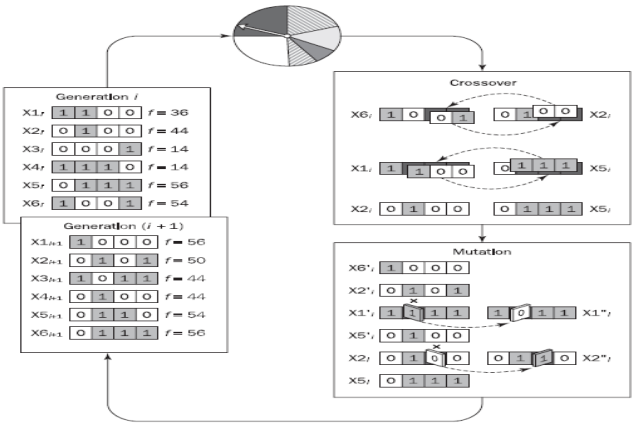
حيث Parent1 و Parent2 هما كروموسومين أبوين، و Offspring1 و Offspring2 هما الكروموسومات النسل، و ExchangePoint هو نقطة التبادل العشوائية. ويعتبر التحور (Mutation) المسؤول عن تغيير قيمة جين محدد في كروموسوم فردي. عادةً ما يتم تطبيق التحور بنسبة منخفضة على الكروموسومات لإدخال تغيير عشوائي في الجينات وتعزيز التنوع الوراثي.يمكن تمثيل هذا العملية بالصيغة الرياضية التالية [7]:

$$MutatedOffspring = Offspring[: MutationPoint] + RandomGene + Offspring[MutationPoint + 1:] \quad (3)$$

حيث (Offspring) هو الكروموسوم الفردي و (MutatedOffspring) هو الكروموسوم النسل المحول و (MutationPoint) هو نقطة التحور العشوائية و (RandomGene) هو جين عشوائي يتم اختياره للتحور.

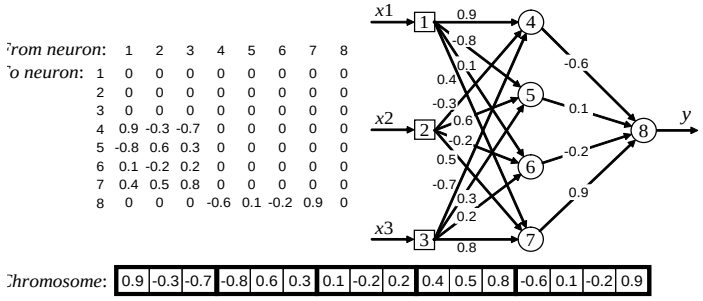
الخطوة 5: إعادة التكرار (Iteration) يتم تكرار الخوارزمية لعدد محدد من الأجيال (Generations) أو حتى تحقيق معيار الإنهاء المحدد. في كل تكرار، يتم تقييم اللياقة والاختيار والتكاثر لإنتاج أجيال جديدة من الكروموسومات.

الخطوة 6: شروط الإنهاء (Termination Criteria) تحدد متى يتوقف تنفيذ الخوارزمية. يمكن أن تكون شروط الإنهاء مثلاً عدد معين من الأجيال، تحقيق قيمة لياقة معينة، انتهاء الوقت المحدد، إلخ. في الشكل (4) توضيح لدورة الحوارزمية الجينية.



الشكل 4: دورة الخوارزمية الجينية [6].

عند استخدام الخوارزميات الجينية يتم تمثيل مجال المشكلة كصبغة (كروموسوم). في الشبكات العصبية يتم البحث عن مجموعة أوزان مثلى للشبكة العصبية المتعددة الطبقات ذات التغذية الأمامية كما في الشكل (5).

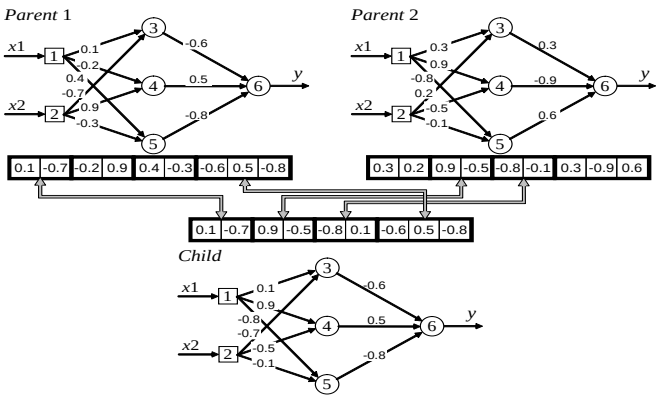


الشكل 5: تشفير مجموعة من الأوزان في الكروموسوم [6].

كخطوة أولى يتم اختيار الأوزان الأولية في الشبكة بشكل عشوائي ضمن نطاق صغير، مثل ½إلى 1. يمكن تمثيل مجموعة الأوزان بواسطة مصفوفة مربعة حيث تمثل القيمة 0 عدم وجود اتصال بين العصبتين.

بشكل إجمالي، هناك 16 رابطًا يحمل وزنًا بين العصبات في الشكل (5). نظرًا لأن الصبغة هي مجموعة من الجينات، يمكن تمثيل مجموعة الأوزان بواسطة صبغة تحتوي على 16 جينًا، حيث يتوافق كل جين مع رابط وزن واحد في الشبكة. بالتالي، إذا قمنا بربط صفوف المصفوفة معًا، واهمال القيم الصفرية سنحصل على صبغة.

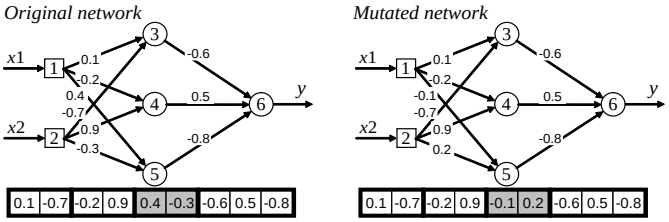
بالإضافة إلى ذلك، يمثل كل صف مجموعة من كافة الروابط المرتبطة بالأوزان الواردة إلى عقدة عصبية واحدة. يمكن اعتبار هذه المجموعة ككتلة وظيفية للشبكة، وبالتالي يجب أن يسمح لها بالبقاء معًا وتمرير المواد الجينية من جيل إلى آخر. لتحقيق ذلك، يجب أن يرتبط كل جين ليس بوزن واحد فحسب، ولكن بمجموعة من جميع الأوزان الواردة إلى عقدة عصبية معينة، كما هو موضح في الشكل (6).



الشكل 6: الكروموسوم في تحسين الوزن [6].

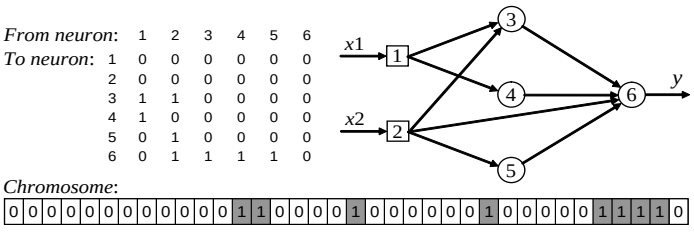
الخطوة الثانية هي تعريف وظيفة اللياقة (fitness function) لتقييم أداء الصبغة (chromosome's performance). يجب على هذه الوظيفة تقدير أداء الشبكة العصبية المعطاة. يمكننا استخدام هنا وظيفة بسيطة إلى حد ما تعرف بمقلوب مجموع مربعات الأخطاء. لتقييم لياقة صبغة معينة، يتم تعيين كل وزن موجود في الصبغة للرابط المقابل في الشبكة. ثم يتم تقديم مجموعة التدريب من الأمثلة إلى الشبكة، ويتم حساب مجموع مربعات الأخطاء. كلما كانت المجموعة أصغر، كانت الصبغة أكثر لياقة. بعبارة أخرى، تحاول الخوارزمية الجينية العثور على مجموعة أوزان تقلل من مجموع مربعات الأخطاء.

الخطوة الثالثة هي اختيار عوامل الجينات - التقاطع والطفرة. يقوم عامل التقاطع بأخذ صبغتين أبويتين وينشئ طفل واحد يحتوي على المواد الجينية من كلا الأبوين. يتم تمثيل كل جين في صبغة الطفل بالجين المقابل في الأبوين المحددين بشكل عشوائي. يقوم عامل الطفرة بتحديد جين عشوائي في صبغة ويضيف قيمة عشوائية صغيرة بين -1 و 1 إلى كل وزن في هذا الجين. يوضح الشكل (7) مثالًا على الطفرة.



الشكل 7: التزاوج او الطفرة في الوزن الأمثل [6].

لتطبيق الخوارزمية الجينية يتعين علينا تحديد حجم السكان وهو عدد الشبكات التي سيتم معالجتها في كل جيل، واحتمالات التقاطع والطفرة، وعدد الأجيال. تحديد حجم السكان يعني تحديد العدد المناسب من الشبكات التي ستساهم في التنوع والاستكشاف في الخوارزمية، ولكنه يؤثر أيضًا على وقت التنفيذ والموارد المطلوبة. بالنسبة لاحتمالات التقاطع والطفرة، فهي تحدد الاحتمالية التي يتم بها تطبيق عمليات التقاطع والطفرة على الصبغات. التقاطع يجمع المواد الجينية من صبغتين أو أكثر لإنشاء صبغة جديدة، في حين تقوم الطفرة بتعديل عشوائي في صبغة واحدة. يجب ضبط هذه الاحتمالات بعناية لتحقيق التوازن بين الاستكشاف والاستغلال في عملية التحسين. عدد الأجيال هو عدد المرات التي يتم فيها تكرار العملية الكاملة من التقاطع والطفرة. زيادة عدد الأجيال يزيد من فرص الحصول على حلول أفضل وأكثر تحسينًا، ولكنه يؤثر أيضًا على وقت التنفيذ.



الشكل 8: ترميز مباشر لطوبولوجيا الشبكة [6].

بناءً على مجموعة من أمثلة التدريب وتمثيل ثنائي لهيكليات شبكة ممكنة، يمكن وصف خوارزمية جينية أساسية بالخطوات التالية:

الخطوة 1: اختيار حجم سكان الصبغات، واحتمالات التقاطع والطفرة، وتحديد عدد حقب التدريب .

الخطوة 2: تعريف وظيفة اللياقة لقياس أداء الصبغة الفردية. عمومًا، يجب أن تعتمد لياقة الشبكة على دقتها وسرعة تعلمها وحجمها وتعقيدها. ومع ذلك، يعتبر أداء الشبكة أهم بكثير من حجمها، وبالتالي يمكن تعريف وظيفة اللياقة بمقلوب مجموع مربعات الأخطاء .

الخطوة 3: توليد عشوائي لسكان صبغة أولية .

الخطوة 4: فك ترميز صبغة فردية إلى شبكة عصبية. نظرًا لأن الشبكة مقيدة بأن تكون ذات تغذية أمامية، يتم تجاهل جميع الروابط الراجعة المحددة في

الصبغة. ثم يتم تعيين الأوزان الأولية للشبكة بأعداد عشوائية صغيرة، مثلاً في النطاق ½إلى 1. ثم تدريب الشبكة على مجموعة تدريب من الأمثلة لعدد معين من حقب التدريب باستخدام خوارزمية الانتشار العكسي. يتم حساب مجموع مربعات الأخطاء وتحديد لياقة الشبكة .

الخطوة 5: تكرار الخطوة 4 حتى يتم النظر في جميع الأفراد في السكان .

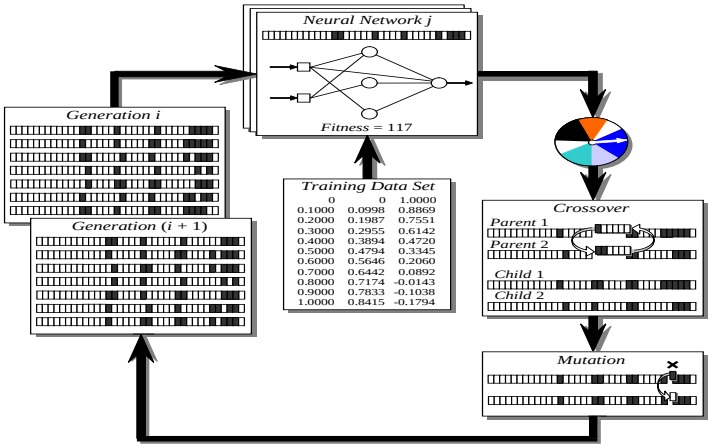
الخطوة 6: اختيار زوجي صبغات للتزاوج، باحتمال متناسب مع لياقتهما .

الخطوة 7: إنشاء زوج من صبغات النسل عن طريق تطبيق عوامل التقاطع والطفرة الجينية. يقوم عامل التقاطع باختيار فهرس صف عشوائي وببساطة يقوم بتبديل الصفوف المقابلة بين الأبوين، مما يؤدي إلى إنشاء طفلين. يقوم عامل الطفرة بتغيير بت واحد أو اثنين في الصبغة بعد احتمال منخفض، مثلاً 0.005.

الخطوة 8: وضع صبغات النسل المُنشأة في السكان الجدد.

الخطوة 9: تكرار الخطوة 6 حتى يصبح حجم سكان الصبغات الجديدة متساويًا مع حجم سكان الصبغات الأولية، ثم استبدال سكان الصبغات الأولية (الأباء) بسكان الصبغات الجديدة (النسل).

الخطوة 10: الانتقال إلى الخطوة 4 وتكرار العملية حتى يتم النظر في عدد محدد من الأجيال.



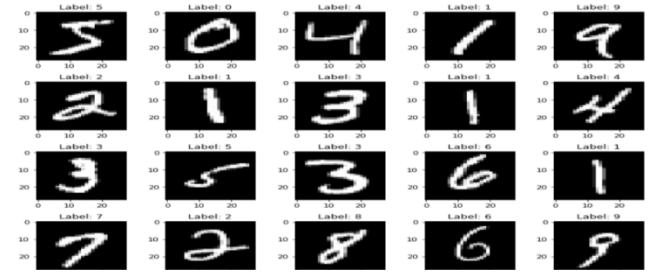
الشكل 9: دورة تطوير طوبولوجيا الشبكة العصبية [6].

مجموعة البيانات:

في هذا البحث تم استخدام مجموعتي بيانات هما:

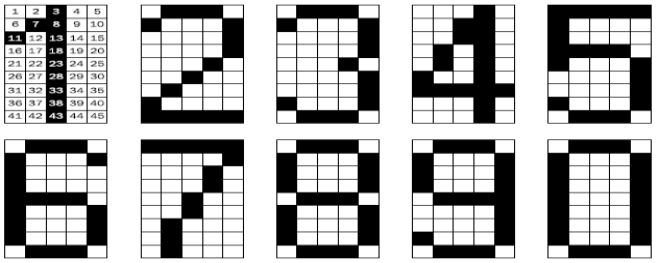
1. قاعدة بيانات MNIST: هي مجموعة بيانات من YANN LECUN'S WEBSITE معيارية معروفة في مجال التعلم الآلي ورؤية الحاسوب. تحتوي على مجموعة كبيرة من الأرقام المكتوبة بخط اليد، وتستخدم عادة لتدريب النماذج التي تهدف إلى التعرف على الأشكال والأنماط في الصور وتقييم أدائها. تتكون مجموعة بيانات MNIST من 60,000 صورة تدريب و 10,000 صورة اختبار، كل صورة ذات تدرج رمادي بحجم 28 × 28 بكسل. تم توسيط الأرقام الموجودة في الصور وتوحيدها، مما يجعلها مناسبة للاستخدام في مجموعة واسعة من التطبيقات [8].

وفي الشكل (9) عينة من مجموعة بيانات MNIST ماخوذة من مجموعة البيانات التي تم العمل عليها لتدريب الشبكة العصبية الاصطناعية. كل صورة ذات حجم 28 * 28 بكسل وقناة لونية واحدة على اعتبار الصور ثنائية اللون (BINARY IMAGE).



الشكل 10: عينة من مجموعة بيانات MNIST.

تم تمثيل كل رقم بخريطة بت 28 * 28 بكسل. في التطبيقات التجارية تستخدم خرائط بت 16 * 16 على الأقل لتعطي ثبات اكبر. في الشكل (5) يبين خرائط بت 9 * 5 صغيرة نسبيا كمثال لخرائط البت لتمييز الرموز. يتم تقسيم الصورة إلى مئات من البيكسل (PIXEL) في البوصة الواحدة، وتمثيل كل بيكسل اما بالرقم 1 (إذا كانت سوداء)، أو بالرقم 0 (إذا كانت بيضاء) وتسمى مصفوفة النقاط الناتجة خريطة البت BIT MAP.



الشكل 11: خرائط البت لتمييز الرموز [6].

2. مجموعة بيانات زهرة اللوتس (Iris):

وهي مجموعة بيانات من KAGGLE متعددة المتغيرات SET DATA MULTIVARIATE، مجموعة البيانات عبارة عن 151 مثال (عينة) لثلاثة أنواع من زهرة IRIS هي (IRIS SETOSA, IRIS VIRGINICA, IRIS VERSICOLOR) تم قياس أربع سمات من كل نوع: طول وعرض SEPALS وطول وعرض PETALS . ومقسمة لكل نوع 50 عينة. وتكون مدخلات الخوارزمية هي طول سيبال(LENGTH_SEPAL) بالسنتيمتر، عرض سيبال (WIDTH_SEPAL)بالسنتيمتر، طول البتلة (LENGTH_PETAL) بالسنتيمتر وعرض البتلة (WIDHT_PETAL) بالسنتيمتر .وتكون الفئة أما SETOSA، VERSICOLOR، أو VIRGINICA وتستخدم كهدف [10] .

النتائج

ركزت هذه الدراسة حول استخدام الخوارزميات الجينية لضبط وتحسين المعاملات الفائقة لشبكات العصبية. تم تطبيقها على مجموعتي بيانات زهرة اللوتس والأرقام المكتوبة بخط اليد باستخدام شبكة عصبية من نوع MLP (MULTILAYER PERCEPTRON).

تم استخدام الخوارزمية الجينية لتحسين المعاملات الفائقة للشبكة العصبية، وهي إحدى الطرق التطورية في مجال الحوسبة الطبيعية. تم تحسين المعاملات الفائقة للشبكة مثل عدد الطبقات الخفية وعدد الدورات ومعدل التعلم وخوارزمية التحسين.

تم إعطاء قيم افتراضية (مبدئية) للنموذجين بالشكل التالي:

حجم الطبقات الخفية = رقم عشوائي من الطبقات الخفية يتراوح عدد الطبقات بين 1 الي 5 طبقات وعدد الوحدات في كل طبقة يتراوح بين 1 الي 10 وحدات، عدد الدورات (الحقب) = قيمة عشوائية بين 100 و 1000 لعدد

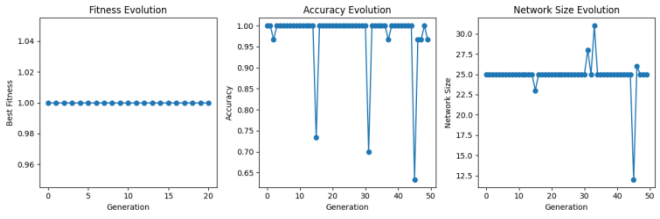
الدورات القصوى للتدريب، معدل التعلم = قيمة عشوائية بين 10^{-1} و 10^{-6}، خوارزمية التحسين = اختيار خوارزمية بشكل عشوائي من بين ثلاثة خوارزميات هي 'SGD' أو 'ADAM' أو 'LBFGS'.

وللخوارزمية الجينية كانت حجم السكان = 50، عدد الأجيال = 20، احتمالية التطور = 0.1، احتمالية التقاطع =0.9، وتم الحصول على النتائج حسب الجدول (1) وهي كالتالي:

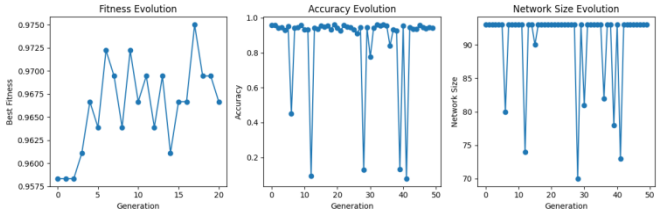
جدول 2: نتائج خوارزمية MPL المحسنة.

مجموعة البيانات	عدد الطبقات الخفية	عدد الدورات	معامل التعلم	خوارزمية التحسين	الدقة
زهرة اللوتس	طبقة واحدة (9 وحدات)	986	0.0029898116264366658	lbfgs	%100
الأرقام المكتوبة بخط اليد	طبقتين (9,10)	559	2.1589793250366946e-06	adam	%94

بناءً على نتائج الخوارزميات الجينية وتحسين المعاملات الفائقة لشبكات العصبية، ومن خلال الشكلين (13,12) يلاحظ أن اللياقة تتحسن مع تقدم الأجيال في الشبكة الاولى حيث تزيد قيمة اللياقة مع كل جيل، يشير ذلك إلى أن الخوارزمية الجينية تنجح في تحسين أداء الشبكة العصبية عبر الأجيال المتعاقبة. حيث يتضح أنه مع كل جيل جديد يتم اكتشاف حلول أفضل وتحسين الأداء بشكل عام، مع وجود انخفاض اللياقة في الشبكة الثانية مما يدل على وجود بعض التحديات التي تواجه الخوارزمية الجينية. يلاحظ أيضا أن معيار الدقة (Accuracy) تزداد مع تقدم الأجيال، حيث تصبح النماذج أكثر قدرة على تصنيف البيانات بشكل ادق، ويصبح الاختيار الجيني وعمليات التكاثر والتحول التي تتم في كل جيل تؤدي إلى تطوير نماذج أكثر دقة. كما يلاحظ أن حجم الشبكة يتغير مع تقدم الأجيال، حيث يزيد أو يقل حسب الخيارات الجينية المطبقة مما يشير الي أن الخوارزمية الجينية تتكيف وتستكشف مجال الحلول بشكل شامل حيث أن الخوارزمية تتعلم تحسين حجم الشبكة في بعض الأحيان لتحقيق أداء أفضل، وفي أحيان أخرى يتم تقليل حجم الشبكة لتحسين الكفاءة.



الشكل 12: دالة اللياقة ومعيار الدقة وحجم الشبكة لبيانات Iris.



الشكل 13: دالة اللياقة ومعيار الدقة وحجم الشبكة لبيانات MNIST.

المناقشة

من خلال مقارنة نتائج الدراسة الحالية مع الدراسات السابقة، يمكن أن نلاحظ بعض النقاط المشتركة والتحسينات التي قدمتها الدراسة. ففي دراسة تشين لي وآخرين (2022)، تم استخدام الخوارزمية الجينية لتحسين معاملات نماذج التعلم العميق المستندة إلى نقل التعلم، وأظهرت النتائج تحسينًا كبيرًا في معيار الدقة. وهذا يتماشى مع نتائج الدراسة الحالية التي أظهرت نسبة دقة تصل إلى 100% لمجموعة بيانات زهرة اللوتس. وفي دراسة جي هون هان وآخرين (2020)، تم استخدام الخوارزمية الجينية لتحسين معاملات شبكات التعلم العميق، وأظهرت النتائج تحسينًا في معدل الدقة ووقت التدريب. وهذا يتماشى مع الدراسة الحالية التي أظهرت تحسينًا في معدلات الدقة وعدد الدورات لمجموعة بيانات الأرقام المكتوبة بخط اليد. بالإضافة إلى ذلك، في دراسة سانغهيوب لي وآخرين (2021)، تم استخدام الخوارزمية الجينية لتحسين هيكل الشبكة العصبية التلافيفية في تشخيص مرض الزهايمر. أظهرت النتائج تحسينًا في مهمة التصنيف وثباتًا أكبر للشبكة المحسنة. على عكس الدراسة الحالية التي كانت غير مستقرة الي حدا ما.

هناك عدة أسباب محتملة لانخفاض دالة اللياقة في خوارزمية الجينات المستخدمة لتحسين المعاملات الفائقة في الشبكة العصبية من نوع MLP وتحديدا على مجموعة بيانات الأرقام المكتوبة بخط اليد. من بين هذه الأسباب:

1. التقصير الزمني: قد يحتاج الخوارزمية الجينية إلى وقت أطول للوصول إلى أفضل الحلول الممكنة. إذا لم تتمكن الخوارزمية من استكشاف مجموعة كافية من الحلول المحتملة قبل انتهاء الوقت المحدد، فقد يؤدي ذلك إلى انخفاض في دالة اللياقة.
2. حجم الشبكة العصبية: قد يكون حجم الشبكة العصبية (عدد الطبقات وعدد الوحدات في كل طبقة) غير مناسب لمجموعة البيانات المستخدمة. إذا كانت الشبكة صغيرة جدًا، فقد يكون لديها قدرة محدودة على استيعاب تعقيد البيانات وقد لا تكون قادرة على تحقيق دقة عالية. وعلى الجانب الآخر، إذا كانت الشبكة كبيرة جدًا، فقد يكون لها تمثيل زائد وتعاني من مشكلة الزيادة في التحجيم.
3. استراتيجية التحسين: قد تحتاج الخوارزمية الجينية إلى ضبط بعض المعاملات والاستراتيجيات المستخدمة في عملية التحسين. على سبيل المثال، يمكن أن تتأثر النتائج بتعيين قيم غير مناسبة لمعدل التحوير وحجم السكان واحتمالية التبديل وغيرها من المعاملات المستخدمة في الخوارزمية الجينية.
4. قيود البيانات: مجموعة البيانات المستخدمة يمكن أن تكون محدودة أو غير متوازنة بطريقة ما، مما يؤثر على أداء الشبكة العصبية.

وقد واجهت الدراسة بعض التحديات التي تؤثر على أداء الخوارزمية الجينية المستخدمة في تحسين المعاملات الفائقة للشبكة العصبية. تشمل ما يلي:

1. الحوسبة الضعيفة: استخدام الحوسبة الرسومية المجانية والمحدودة تؤثر على سرعة تنفيذ الخوارزمية. قد يكون الوقت المطلوب للحصول على نتائج

دقيقة أطول من المتوقع بسبب ذلك.

2. عدم استقرار شبكة الإنترنت: فقدان الاتصال أثناء تنفيذ الخوارزمية تؤثر سلبا على استمرارية التحسين.

3. إهمال وقت التنفيذ: بسبب الحوسبة الضعيفة و عدم استقرار شبكة الإنترنت تم اهمال وقت التنفيذ ولهذا السبب يحتمل وجود صعوبة في استكمال التحسين بشكل كامل أو تجربة مختلفة للمعاملات الفائقة.

الخاتمة

في هذه الدراسة، تم استخدام الخوارزمية الجينية لضبط وتحسين المعاملات الفائقة لشبكات العصبية، وتم تطبيقها على شبكة العصبية MLP باستخدام مجموعتي بيانات زهرة اللوتس والأرقام المكتوبة بخط اليد. توصلت الدراسة إلى أن الخوارزمية الجينية تلعب دورًا حاسمًا في تحسين المعاملات الفائقة للشبكات العصبية وزيادة دقة التصنيف. من خلال تحسين المعاملات الفائقة للشبكة العصبية MLP باستخدام الخوارزمية الجينية، تم تحقيق نتائج جيدة. حيث تم تحقيق دقة تصل إلى 100% لمجموعة بيانات زهرة اللوتس، مما يدل على أداء ممتاز للتصنيف. وبالنسبة لمجموعة بيانات الأرقام المكتوبة بخط اليد، تم الحصول على دقة تصل إلى 94%، مما يشير إلى أداء جيد وفعالية النموذج في التعرف على الأرقام. تؤكد هذه النتائج أهمية وفاعلية الخوارزمية الجينية في تحسين ضبط المعاملات الفائقة للشبكات العصبية، مما يمكن من تحسين أداء النماذج وزيادة قدرتها على التعرف على الأنماط والتصنيف بدقة. كما تشير النتائج إلى أن استخدام الخوارزمية الجينية يمكن أن يؤدي إلى تحسين معدلات الدقة وعدد الدورات المطلوبة للتدريب، مما يعزز كفاءة عملية التعلم. ومع ذلك، تحدثت الدراسة عن بعض التحديات التي يمكن مواجهتها في تطبيق الخوارزمية الجينية لتحسين المعاملات الفائقة للشبكات العصبية. من بين هذه التحديات هي الحوسبة الضعيفة وعدم استقرار شبكة الإنترنت، والتي يمكن أن تؤثر على سرعة التحسين واستمراريته. بالإضافة إلى ذلك، قد يكون هناك تحديات في اختيار قيم المعاملات واستراتيجيات التحسين التي تؤثر على أداء الخوارزمية الجينية. في المستقبل، يمكن استكشاف تطبيقات أخرى للخوارزمية الجينية في تحسين المعاملات الفائقة لشبكات العصبية وتوسيع نطاق الدراسات الحالية. يمكن أيضًا تحسين الأداء من خلال استكشاف مجموعات بيانات أكبر وتعديل قيم المعاملات واستراتيجيات التحسين لتحقيق أداء أفضل. باختصار، تظهر النتائج أن الخوارزمية الجينية تلعب دورًا مهمًا في تحسين المعاملات الفائقة للشبكات العصبية وتحسين أدائها في التصنيف. تعد هذه الدراسة خطوة مهمة في تطوير الأدوات والتقنيات المتقدمة في مجال الذكاء الاصطناعي وتعزيز فهمنا لكيفية تحسين أداء الشبكات العصبية.

الاعمال المستقبلية:

1. تطبيق الخوارزميات الجينية على مجموعات بيانات مختلفة: يمكنك استكشاف تأثير استخدام خوارزميات الجينات في تحسين المعاملات الفائقة على مجموعات بيانات أخرى. يمكنك اختيار مجموعات بيانات متنوعة وتحليل كيفية تأثير الخوارزميات الجينية على أداء الشبكة العصبية على هذه المجموعات.

2. مقارنة بين خوارزميات التحسين: يمكنك دراسة ومقارنة بين خوارزميات التحسين المختلفة للشبكات العصبية. يمكنك استكشاف تأثير استخدام خوارزميات مختلفة على أداء الشبكة العصبية ومعاملاتها الفائقة على مجموعة بيانات محددة.

قائمة المراجع

[1]- Chen Li, JinZhe Jiang, YaQian Zhao, RenGang Li, EnDong Wang, Xin Zhang, and Kun Zhao "Genetic algorithm based hyper-parameters optimization for transfer convolutional neural network", Proc. SPIE 12285, International Conference on Advanced Algorithms and Neural Networks (AANN 2022), 1228512 (15 June 2022); https://doi.org/10.1117/12.2637170

[2]- Han, Ji-Hoon; Choi, Dong-Jin; Park, Sang-Uk; Hong, Sun-Ki (2020). "Hyperparameter Optimization Using a Genetic Algorithm Considering Verification Time in a Convolutional Neural Network." Journal of Electrical Engineering & Technology. doi:10.1007/s42835-020-00343-7.

[3]- Erden, C. Genetic algorithm-based hyperparameter optimization of deep learning models for PM2.5 time-series prediction. Int. J. Environ. Sci. Technol. 20, 2959–2982 (2023). https://doi.org/10.1007/s13762-023-04763-6

[4]- Lee, S.; Kim, J.; Kang, H.; Kang, D.-Y.; Park, J. Genetic Algorithm Based Deep Learning Neural Network Structure andHyperparameter Optimization. Appl. Sci. 2021, 11, 744. https://doi.org/ 10.3390/app11020744

[5]- A.E. Eiben and J.E. Smith, "Introduction to Evolutionary Computing, Second Edition," Natural Computing Series, ISSN 1619-7127, doi: 10.1007/978-3-662-44874-8, 2015.

[6]- Negnevitsky, M. (2011). Artificial Intelligence: A Guide to Intelligent Systems (3nd). Pearson Education Limited.

[7]- Sushruta Mishra, Soumya Sahoo, Mamata Das, "Genetic Algorithm: An Efficient Tool for Global Optimization," Advances in Computational Sciences and Technology (ACST), Volume 10, Number 8, 2017, pp. 2201-221.

[8]- Y. LeCun, C. Cortes, and C. Burges, "MNIST handwritten digit database," Retrieved from http://yann.lecun.com/exdb/mnist/, Accessed on 27-1-2023.

[9]- Sahad, A. K., & Khowai, H. (2022). A comparison of artificial neural networks and support vector regression models for forecasting oil prices during the period 2000-2022. Journal of Economic and Administrative Research, 61(20), 38-83.

[10]- Iris dataset UCI. (2023). Kaggle. Retrieved on April 11, 2023, from https://www.kaggle.com/datasets/jillanisofttech/iris-dataset-uci.