مجلة جامعة طرطوس للبحوث والدراسات العلمية _ سلسلة العلوم الهندسية المجلد (٩) العدد (٧) ٢٠٢٥
Tartous University Journal for Research and Scientific Studies - engineering Sciences Series Vol. (9) No. (7) 2025

تطوير نظام لكشف وتصنيف الهجمات في الشبكات الحاسوبية بالاعتماد على خوارزمية الخلايا الجذعية المناعية والخوارزمية الجينية

د. يعرب ديوب*

د. جعفر سلمان **

سالي محمد عيسى ***

(تاريخ الإيداع ٢٠٢٥/٤/٩ . قُبِل للنشر في ٢٠٢٥/٧/٦)

□ ملخّص □

تزامناً مع التسارع الكبير في استخدام الانترنت والاعتماد على تقنياته في كافة مجالات الحياة وخاصة مع التطورات التكنولوجية المتلاحقة، أصبحت الشبكات الرقمية بكافة عناصرها عرضه للهجمات السيبرانية المعقدة، فكلما زادت حركة البيانات وتطورت أساليب تبادلها، ارتفعت معها فرص التعرض للاختراق الذي يؤثر على سلامة المعلومات واستمرارية الخدمات، هذا بدوره يتطلب وجود أنظمة ذكية ومتقدمة قادرة على كشف التهديدات والتعامل معها بكفاءة تُضاهي التطور الحاصل في أساليب الهجوم وتنوعها وتضمن استقرار الأنظمة والشبكات وحماية حركة البيانات فيما بينها.

يهدف هذا البحث إلى دراسة دور خوارزمية الخلايا الجذعية المناعية في تطوير أنظمة لكشف وتصنيف الهجمات في الشبكات الحاسوبية، لما تتميز به من القدرة على معالجة البيانات الكبيرة وتحليل الأنماط المتنوعة والمعقدة بفعالية ودقة تجعلها أداة واعدة في مجال الأمن السيبراني والتعامل مع التحديات المتطورة.

قدمنا في هذا البحث نظاماً ذكياً لتحسين كشف وتصنيف التهديدات السيبرانية في الشبكات الحاسوبية بالاعتماد على خوارزمية الخلايا الجذعية والخوارزمية الجينية وتقييم أداءه في بيئات شبكية متنوعة باستخدام قاعدة البيانات UNSW–NB15 المعيارية وبرنامج VSCode مع لغة بايثون بما تضمّه من مكتبات خاصة بالتعلم الآلي وتحليل البيانات.

أظهرت النتائج النهائية لهذه الدراسة أداء مميز للنظام المقترح في تصنيف الهجمات وتحديد النشاط الضار في الشبكة مع إمكانية التعامل مع النشاطات المشبوهة غير المؤكدة وتنصيفها بآلية ذكية تحقق دقة تصنيف عالية وصلت إلى 97% مع معدل إنذارات كاذبة منخفض جداً وصل إلى 4% بالإضافة إلى قابلية التنفيذ في الوقت الفعلي مما يمنح النموذج كفاءة عالية ويبرز الدور الفعال للخوارزميات الحيوية في الأمن السيبراني.

الكلمات المفتاحية: أنظمة الكشف عن التسلل، التصنيف، حركة البيانات في الشبكات، أنظمة المناعة الاصطناعية، خوارزمية الخلايا الجذعية، قاعدة البيانات UNSW–NB15، الخوارزمية الجينية، الدقة، معدل الإنذارات الكاذبة.

* أستاذ في قسم هندسة تكنولوجيا المعلومات – كليّة هندسة تكنولوجيا المعلومات والاتصالات – جامعة طرطوس–سوريا.

** أستاذ مساعد في قسم هندسة تكنولوجيا المعلومات – كليّة هندسة تكنولوجيا المعلومات والاتصالات – جامعة طرطوس–سوريا.

*** طالبة دكتوراه – قسم هندسة تكنولوجيا المعلومات – كليّة هندسة تكنولوجيا المعلومات والاتصالات – جامعة طرطوس–سوريا.

مجلة جامعة طرطوس للبحوث والدراسات العلمية _ سلسلة العلوم الهندسية المجلد (٩) العدد (٧) ٢٠٢٥
Tartous University Journal for Research and Scientific Studies - engineering Sciences Series Vol. (9) No. (7) 2025

Development of a System for Detecting and Classifying Attacks in Computer Networks based on the Dendritic Cell Immune Algorithm and the Genetic Algorithm

Dr. Yaroub Dayoub*
Dr. Jaafar Salman **
Sally Mohammad Issa ***

(Received 9/4/2025 . Accepted 6/7/2025)

□ ABSTRACT □

With the growing reliance on its technologies across all aspects of life particularly in light of continuous technological advancements digital networks and their components have become increasingly vulnerable to sophisticated cyber-attacks. As data traffic intensifies and communication methods evolve, the risk of breaches that compromise information integrity and service continuity also rises. This necessitates the development of intelligent and advanced systems capable of detecting threats and responding to them efficiently, matching the pace and complexity of modern attack strategies to ensure system stability and secure data exchange.

This research aims to investigate the role of the immune-inspired Stem Cell Algorithm in enhancing intrusion detection and classification systems within computer networks, due to its ability to process large-scale data and analyze diverse, complex patterns with high accuracy making it a promising tool in cybersecurity for addressing evolving threats.

In this research, we propose an intelligent system that enhances the detection and classification of cyber threats in computer networks by integrating the Stem Cell Algorithm with Genetic Algorithms. The system's performance was evaluated in diverse network environments using the benchmark UNSW-NB15 dataset and implemented via the VSCode environment using Python and its relevant machine learning and data analysis libraries.

The final results of this study demonstrated the effectiveness of the proposed system in accurately classifying attacks and identifying malicious activity in the network, with the ability to intelligently handle uncertain suspicious behavior. The model achieved a high classification accuracy of 97% and a very low false alarm rate of 4%, along with real-time execution capability, highlighting its high efficiency and the significant role of bio-inspired algorithms in enhancing cybersecurity.

Keywords: Intrusion detection systems, classification, network traffic, artificial immune systems, Dendritic Cell Algorithm, UNSW-NB15 database, genetic algorithm, accuracy, FAR.

*Professor, Information Technology Engineering Department, Information and communication Technology Engineering, Tartous University, Syria.
**Assistant Professor, Information Technology Engineering Department, Information and communication Technology Engineering, Tartous University, Syria.
***PhD Student, Information Technology Engineering Department, Information and communication Technology Engineering, Tartous University, Syria.

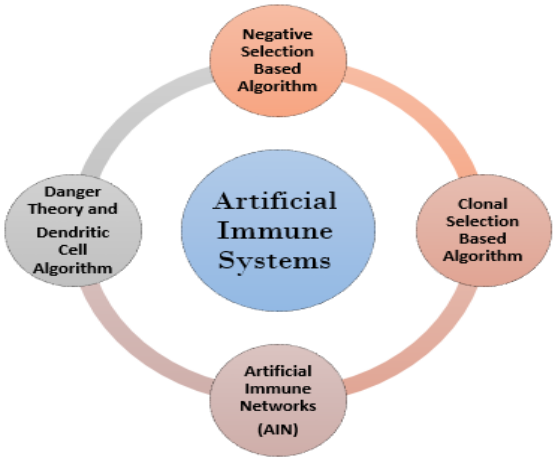
١. مقدمــة:

لعبت أنظمة المناعة الاصطناعية Artificial Immune System (AIS) دوراً فعالاً في العديد من تطبيقات العالم الحقيقي منذ نشأتها في تسعينيات القرن العشرين [1] وحتى وقتنا هذا بوصفها واحدة من أحدث تقنيات الذكاء الحاسوبي Computational Intelligence المستوحاة من الطبيعة وبخاصة علم المناعة البيولوجي وأنظمته الحيوية ذات المواصفات المميزة والقدرات الفعالة التي تجعلها جذابة للاستثمار في مجال الحوسبة بغرض العمل على توفير بيئة حماية تحمل خصائص دفاعية مستندة إلى نماذج جهاز المناعة البشري.

عمل الباحثون بالتعاون مع علماء البيولوجيا على تطوير خوارزميات مناعية تعتمد في مبادئها وأهدافها على نظام المناعة الحيوي لتكون قادرة على التعلم، التكيف، واكتشاف أنماط البيانات. كما استمرت هذه الأبحاث في التطور بالتوازي مع كل تطور يُكتشف في المناعة الأساسية وآلياتها الدفاعية بغرض حل مشاكل الأمثلة والتحسين [2]، انطلاقاً من خوارزمية الاختيار السلبي(NSA) Negative Selection Algorithm [3] التي استُخدمت على نطاق واسع في كشف الشذوذ والتهديدات غير المعروفة، وصولاً إلى خوارزمية الاختيار النسلي Clonal Selection Algorithm (CSA) المعتمدة على مبدأ التطور المناعي لتحسين الحلول بناءً على التكاثر والطفرات.

إلا أنّ خوارزميات الجيل الأول من أنظمة المناعة كانت تعاني من مشاكل تتعلق بإمكانية التوسع والعدد الكبير من الإنذارات الكاذبة مما أدى لنتائج ضعيفة من حيث الدقة والأداء. هذا بدوره دفع الباحثين إلى تطوير خوارزميات تحاكي آليات بيولوجية متقدمة وحديثة تتصف بالمتانة، الدقة والتنظيم الذاتي تحت ما يسمى بـ "خوارزميات الجيل الثاني" والتي كان أهمها خوارزمية الخلايا الجذعية Dendritic Cell Algorithm (DCA) التي تحاكي آلية عمل الخلايا الجذعية ومبادئ نظرية الخطر Danger Theory في جهاز المناعة الحيوي، والمستخدمة في العديد من تطبيقات الحوسبة المتعلقة بالأمن حيث أظهرت نتائج فعالة ودقيقة[1] .

يوضح الشكل (1) الخوارزميات المدرجة ضمن أنظمة المناعة الاصطناعية والتي تحاكي كل منها في آلية عملها وظيفة من وظائف جهاز المناعة الحيوي.



الشكل (1): أنظمة المناعة الاصطناعية وخوارزمياتها

٢. هدف البحث:

يهدف هذا البحث إلى تطوير نظام لكشف التسلل في الشبكات Intrusion Detection System (IDS) بالاعتماد على خوارزمية الخلايا الجذعية المناعية (DCA) والخوارزمية الجينية وذلك من خلال محاكاة آلياتها في الاختيار والكشف عن التهديدات غير المعروفة والتمييز بين سلوك حركة البيانات (Data Traffic) الطبيعي والضار في الشبكة.

نسعى من خلال هذا البحث إلى العمل على تحسين دقة وكفاءة نظام الكشف عن التسلل من خلال الاستفادة من قدرة خوارزمية DCA في معالجة البيانات بطريقة غير خطية وبالتالي القدرة على التعرف على الأنماط والتكيف مع أي تهديد جديد بالتوازي مع تقليل معدل الإنذارات الكاذبة مما يساهم في توفير بيئة الكترونية أكثر أمان.

٣. مواد وطرق البحث:

اعتمدنا في الدراسة العملية وتحليل النتائج المنبثقة عنها على برنامج (VScode) وهو أداة مجانية مفتوحة المصدر تم تطويرها بواسطة شركة مايكروسوفت يتميز بالسرعة والمرونة، بالإضافة لإمكانية تصحيح الأخطاء ودعم الإضافات كما أنّه يعمل على كافة أنظمة التشغيل.

يدعم (VScode) العديد من لغات البرمجة وخاصة لغة (Python) التي قمنا باستخدامها كأداة رئيسية لمعالجة البيانات وتحليلها وبناء الكود البرمجي للخوارزميات المستخدمة حيث أثبتت فعالية ومرونة عالية في برمجة تطبيقات الذكاء الصنعي لما توفره من مكتبات قوية تدعم العمليات الحسابية ومعالجة البيانات مثل (Tensorflow, Pandas, Sklearn, NumPy).

٤. نظام كشف التسلل في الشبكات Intrusion Detection System (N-IDS) in Networks:

إن نظام كشف التسلل IDS عبارة عن جهاز أو تطبيق برمجي يعمل على مراقبة الشبكة أو نشاطات النظام بحثًا عن أي نشاط ضار معروف أو حدث مريب أو انتهاكات لسياسة الأمان من خلال استغلال للثغرات الأمنية ويقوم بإرسال التنبيهات إلى مسؤولي النظام في الوقت الفعلي، فهو يعمل كطبقة دفاعية إضافية لمساعدة مسؤولي الأمن في التعرف على التهديدات والاستجابة لها تفادياً لحصول ضرر كبير في الموارد والبيانات [4].

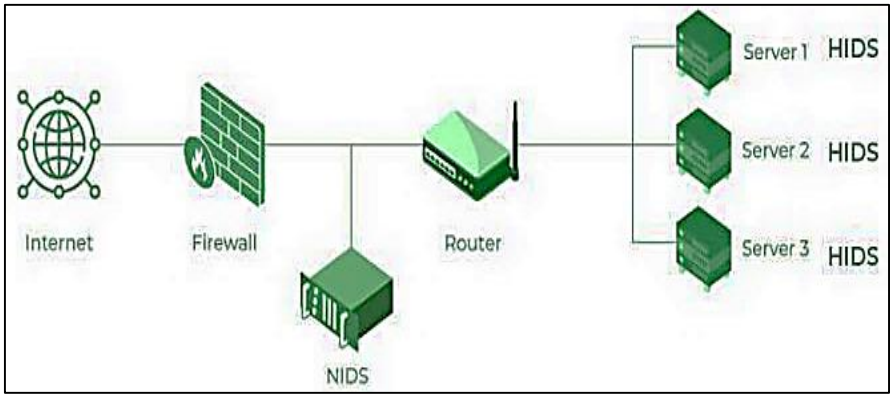
لدينا نوعين مختلفين من أنظمة كشف التسلل وهما:

١- نظام كشف التسلل القائم على الشبكة N-IDS:

يوضع هذا النظام في نقاط معينة على الشبكة بحيث يستطيع مراقبة كل حركة الشبكة traffic من وإلى الأجهزة المتصلة بالشبكة حيث يتضمن قاعدة بيانات للهجمات والتهديدات المعروفة التي يمكن أن تهدد الشبكة ويعمل على تحليل حركة البيانات من وإلى الشبكة، وعند تطابق أي منها مع الموجود ضمن قاعدة البيانات يتم تصنيفها على أنها نشاط مشبوه وارسال انذار لإعلام المعنيين بذلك، كما يمكنه الاعتماد على التوقيع Signature الخاص بالحزم ومقارنته مع توقيعات مشابهة لحزم ضارة معروفة ضمن الشبكة.

٢- نظام كشف التسلل القائم على المضيف H-IDS:

يعمل على مراقبة المضيفين أو الأجهزة في الشبكة من خلال تفقد كل الحزم packets الداخلة والخارجة من وإلى الجهاز وارسال التنبيهات عند ملاحظة نشاطات مشبوهة [5].



الشكل (2): نظام كشف التسلل في الشبكات N-IDS

٥. خوارزمية الخلايا الجذعية المناعية Dendritic Cell Immune Algorithm(DCA):

في الآونة الأخيرة أصبحت الخوارزميات المستوحاة من البيولوجيا والطبيعة البشرية خاصة تلعب دوراً حاسماً في العديد من تطبيقات التكنولوجيا الحديثة بوصفها حل فعال للعديد من المشاكل والتحديات التي تتزايد مع تزايد التطور التقني، وأهمها نظام المناعة الاصطناعية Artificial Immune System (AIS) وخوارزمياته الذكية التي تندرج تحت مسمى الذكاء الحسابي Computational Intelligence (CI) [6].

إحدى أهم هذه الخوارزميات والتي اعتمدنا عليها في انجاز هذا البحث هي خوارزمية الخلايا الجذعية (DCA) والتي تم اقتراحها لأول مرة عام 2006 من قبل الباحثة Julie Tuwai Kim وفريقها في جامعة نوتغهام Nottingham كمحاولة لاستغلال النماذج المناعية البيولوجية ومحاكاة آليات دفاعها الفعالة في مجال أمن المعلومات والشبكات، فهي تلعب دوراً حاسماً في فعالية الجهاز المناعي ضد التهديدات والاستجابة المثلى لها من خلال العمل على تمييز العناصر الخطيرة من الآمنة داخل الجسم.

تم محاكاة آلية عمل الخلايا المناعية في حماية الشبكة من أي تهديد من خلال كشف التسلل والتعرف على الأنماط الخبيثة في حركة بيانات الشبكة والتعامل معها بفعالية ودقة عالية[7]، فهي تتمتع بالديناميكية العالية مما يجعها قادرة على التعرف على النماذج، بالإضافة لقابلية التكيف العالية مع الهجمات الجديدة دون الحاجة إلى إعادة البرمجة.

٥,١ آلية عمل خوارزمية DCA:

تحاكي الخوارزمية الطريقة التي تكتشف بها الخلايا الجذعية التهديدات في الجسم، ثم تنقل هذه المعلومات إلى الخلايا المناعية الأخرى لتنشيط استجابة مناعية، فهي تعالج المشكلات بطريقة مشابهة في الخوارزمية، يتم البحث بكفاءة عن حلول مثالية للمشكلة بطريقة مشابهة لكيفية تعامل الخلايا الجذعية البيولوجية مع المحفزات.

- تتضمن خوارزمية DCA ثلاث عمليات أساسية مشابهة لعمل الخلايا المناعية البيولوجية وهي [8]:

١. **الكشف (Detection):** تعمل الخلايا الجذعية الاصطناعية على تحديد الأنماط الشاذة أو التهديدات في البيئة المدروسة.

٢. **التفاعل (Interaction):** بعد تحديد التهديدات، تبدأ الخلايا بالتفاعل مع خلايا المناعة الأخرى، محاكاةً بذلك عملية نقل الإشارة المناعية بين عناصر الجهاز المناعي.

٣. **الاستجابة (Response):** في النهاية، تعمل هذه الخلايا على توليد استجابة مناعية من خلال معالجة البيانات المتاحة والبحث عن الحل الأمثل من بين مجموعة الحلول.

- تتواصل الخلايا الجذعية ضمن جهاز المناعة البيولوجي وتتفاعل مع بعضها من خلال مجموعة إشارات يتم استخدامها لإيصال معلومات حول حالة الجسم ونوع التهديد المكتشف، حيث تؤدي كل منها دورًا في تحديد جودة الحلول والتفاعل مع المحفزات في البيئة لتوجيه الخلايا الجذعية نحو تحسين الأداء والتصفية المستمرة للحلول غير المثلى وهي [9]:

- **إشارة الخطر Danger Signal (DS)**: هي مؤشر على وجود خطر تنبعث من الخلايا الشجرية عندما تكتشف محفزات ضارة أو تهديدات (مثل البكتيريا أو الفيروسات).
- **اشارات الأمان Safe Signals (SS)**: هي إشارات تُرسل عندما لا يكون هناك تهديد أو خطر وبالتالي البيئة أو المحفز آمن ولا يحتاج إلى استجابة مناعية حادة.
- **الإشارات المساعدةPAMP Signals :** هي إشارات مرتبطة بالأنماط الجزيئية التي تكون مميزة للكائنات الدقيقة المسببة للأمراض (مثل البكتيريا أو الفيروسات)، والتي تُستخدم من قبل الخلايا المناعية لاكتشاف التهديدات وتحفيز الاستجابة المناعية.

يوضح الجدول التالي آلية محاكاة مفاهيم خوارزمية الخلايا الجذعية المناعية ومفرداتها في خوارزمية DCA الاصطناعية:

الجدول (1): مقارنة بين مصطلحات المناعة الحيوية ومقابلاتها في أنظمة المناعة الاصطناعية

المصطلح البيولوجي	المفهوم البيولوجي	المصطلح الاصطناعي	المفهوم ضمن خوارزمية DCA
الخلايا الجذعية Dendritic Cells	خلايا مناعية تكتشف التهديدات (مثل الفيروسات أو البكتيريا).	**الخلايا الاصطناعية Artificial Cells**	مجموعة حلولًا ممكنة يتم تقييمها وتصفية جزء منها وفقًا لجودة وتفاعل كل منها مع بيانات أخرى.
مسببات المرض Antigens	المواد الغريبة (التهديدات) داخل الجسم التي يكتشفها الجهاز المناعي ويستجيب لها	**المحفزات Stimuli**	تمثل المتغيرات التي تحاول الخلايا الجذعية العثور عليها.
الاستجابة المناعية Immune Response	تفاعل يحدث عندما تكتشف الخلايا الجذعية التهديدات وترسل إشارات لتحفيز استجابة مناعية	**تحسين الحلول Optimization**	تبادل المعلومات بين الخلايا الجذعية حول الحلول ومدى فعاليتها.
التصفية المناعية Clonal Selection	عملية تصفية Antigens غير الفعالة.	**التصفية الاصطناعية (حلول أفضل)**	عملية تصفية الحلول غير الملائمة بعد تقييمها بغرض الحفاظ على أفضل الحلول.
الذاكرة المناعية Immunological Memory	يتم تخزين Antigens التي تم التعامل معها سابقًا لتحسين الاستجابة المناعية المستقبلية.	**تخزين الحلول**	تخزين الحلول الجيدة لاستخدامها مستقبلاً في تحسين الحلول بمرور الوقت.
إشارة الخطر Danger Signal	مؤشر على وجود خطر يجب أن يتم التعامل معه بسرعة من قبل جهاز المناعة.	**إشارة الخطر DS**	مؤشر إلى أن الحل يحتوي على خصائص قد تؤدي إلى تحسين الأداء بشكل كبير.

إشارة الأمان Safe Signal	تشير إلى أن البيئة أو المحفز آمن ولا يحتاج إلى استجابة مناعية حادة.	**إشارة الأمان SS**	تشير إلى أن الحل الذي تم التفاعل معه آمن أو مناسب بما يكفي لاستمرار المعالجة أو تحسينه في المستقبل.
إشارة PAMP Signals	تُستخدم من قبل الخلايا المناعية لاكتشاف التهديدات وتحفيز الاستجابة المناعية.	**الإشارة المساعدة PAMP**	تشير إلى أن الحل يحتوي على أنماط مشبوهة لكن ليست ضارة بشكل مؤكد

- تعمل الخلايا الجذعية كوسيط ذكي بين المدخلات (البيانات) والمخرجات (التصنيف) وذلك من خلال تحليل الإشارة التي تتلقاها وتحليل البيانات إلى طبيعية أو ضارة، فبحسب نوع الإشارة الواردة إلى الخلية تتغير حالتها إلى واحدة من الحالات التالية [10]:

١. خلية غير ناضجة iDC(Immature DC): ليس لديها بيانات كافية بعد لاتخاذ أي قرار.
٢. خلية ناضجة mDC (Mature DC) : تشير إلى وجود نشاط ضار وبالتالي تهديد مؤكد.
٣. خلية شبه ناضجة smDC(Semi-mature DC): تشير إلى أن البيانات مشبوهة لكن تحتاج لبيانات أكثر لتؤكد ضررها بدقة.

- بناءً على حالة الخلايا المناعية وباسقاط ذلك على مجال كشف التسلل في الشبكات الذي يمثل محور بحثنا يتم تصنيف تدفقات الشبكة إلى:

١. نشاط طبيعي Benign: إذا تم تصنيف معظم الخلايا المناعية كخلايا غير ناضجة (iDC).
٢. نشاط مشبوه Suspect: إذا تم تصنيف معظم الخلايا المناعية كخلايا شبه ناضجة (smDC).
٣. نشاط ضار Malicious: إذا تم تصنيف معظم الخلايا المناعية كخلايا ناضجة (mDC).

- يتم حساب قيمة إشارة الأمان SS والتي تشير إلى حركة طبيعية داخل الشبكة وفق المعادلة (1) [11]:

$$SS = \frac{1}{N} \sum_{i=1}^{N} w_i * f_i$$

المعادلة[1]

حيث أنّ: N: عدد الميزات المرتبطة بالسلوك الطبيعي.
w_i: الوزن المخصص لكل ميزة (يحدد مدى تأثيرها على الإشارة).
f_i : القيم الطبيعية للميزات.

- يتم حساب قيمة إشارة الخطر DS والتي تمثل مقدار الشذوذ في النشاط الشبكة وفق المعادلة (2)[11] :

$$DS = \frac{1}{M} \sum_{j=1}^{M} w_j * g_j$$

المعادلة[2]

حيث أنّ: M: عدد الميزات التي تشير إلى سلوك غير طبيعي.
w_j: الوزن المخصص لكل ميزة.
g_j : القيم الطبيعية للميزات.

- يتم حساب قيمة الإشارة المساعدة PAMPوالتي تمثل دليلاً قاطعاً على وجود هجوم وفق المعادلة (3)[10]:

$$PAMP = \sum_{k=1}^{k} w_k * h_k$$

المعادلة[3]

حيث أنّ: K: عدد الميزات المرتبطة بالهجمات المؤكدة.
w_k: الوزن المخصص لكل ميزة.
h_j : القيم التي تشير إلى هجوم مؤكد.

- وأخيراً يتم حساب مستوى النضوج للخلية والذي سيتم الاعتماد عليه في تصنيف البيانات إلى سليمة أو مشبوهة أو ضارة وفق المعادلة (4)[10]:

$$\mathrm{MCAV} = \frac{\mathbf{N-attack}}{\mathbf{N-total}}$$

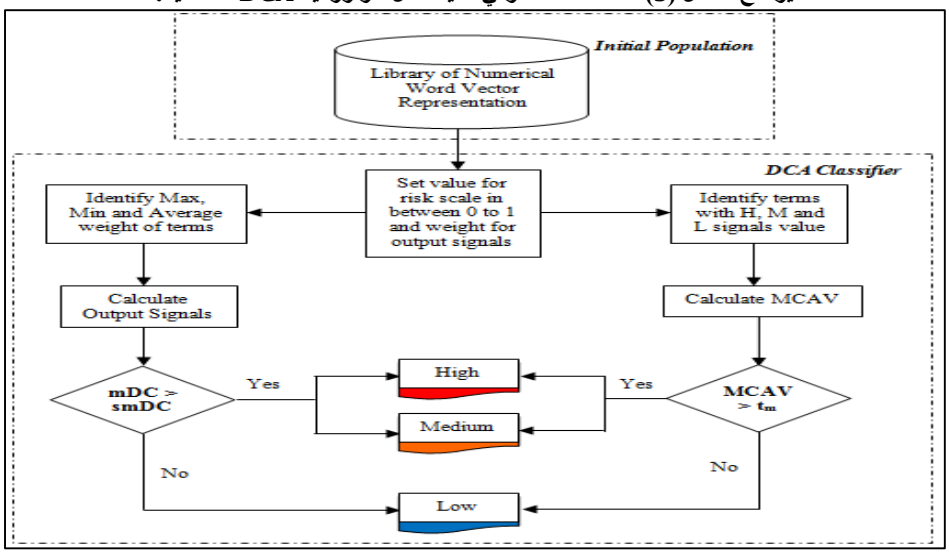
المعادلة[4]

حيث أنّ: N-attack: عدد المرات التي يتم تصنيف الخلية على أنها مشبوهة.
N-total: عدد المرات التي ظهرت فيها الخلية ضمن سياق مجموعة البيانات.

تعتمد الخوارزمية على تراكم قيم الإشارات المختلفة لتحديد التصنيف حيث نميز الحالات التالية:

- إذا كان مستوى النضوج للخلية أصغر من قيمة عتبة الشذوذ الدنيا T_{low} عندها يتم تصنيف البيانات على أنها سليمة.
- إذا كان مستوى النضوج للخلية أكبر من قيمة عتبة الشذوذ العليا T_{high} عندها يتم تصنيف البيانات على أنها ضارة.
- إذا كان مستوى النضوج للخلية بين قيمتي عتبة الشذوذ العليا والدنيا عندها يتم تصنيف البيانات على أنها شبه ضارة.

يوضح الشكل (3) المخطط الصندوقي لآلية عمل خوارزمية *DCA* المناعية:



الشكل (3): آلية عمل خوارزمية DCA المناعية

٥,٢ معايير تقييم أداء الخوارزمية:

➢ **دقة التصنيف classification accuracy**: وهي تمثل قدرة النموذج على توقع الفئة المستهدفة بشكل صحيح وتقاس بالنسبة المئوية وفق المعادلة (5): [12]

$$Accuracy = \frac{TP+TN}{TP+FN+FP+TN}$$

المعادلة [5]

حيث أنّ:

- TP: نسبة البيانات السليمة المصنفة بشكل صحيح.
- TN: نسبة البيانات المشبوهة المصنفة بشكل صحيح.
- FP: نسبة البيانات السليمة المصنفة بشكل خاطئ.
- FN: نسبة البيانات المشبوهة المصنفة بشكل خاطئ.

➢ **مقياس Root Mean Squared Error (RMSE)**: وهو عبارة ن مقياس إحصائي يستخدم لقياس مدى دقة نموذج التصنيف وتقييم جودته من خلال حساب مقدار الانحراف بين القيم الفعلية والقيم المتوقعة بواحدات البيانات الأصلية [13].

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(P_i - \widehat{P}_i)^2}$$

المعادلة [6]

حيث أنّ: P_i: القيم الفعلية (التصنيفات الحقيقية في البيانات).

$\widehat{P}_i$: القيم المتوقعة (تصنيفات نموذج التصنيف).

N: عدد العينات في مجموعة البيانات.

- **معدل اكتشاف الهجمات Detection Rate (DR) [14]:**

$$DR = \frac{TP}{TP+FN}$$

المعادلة [7]

- **معدل الإنذارات الخاطئة False Alarm Rate (FAR) [12]:**

$$FAR = \frac{FP}{FP+TN}$$

العلاقة [8]

٦. توصيف مجموعة البيانات:

في هذه الدراسة، تم الاعتماد على مجموعة البيانات UNSW NB15 وهي عبارة عن مجموعة بيانات حديثة تم تطويرها وجمعها كجزء من مشروع بحثي في جامعة نيو ساوث ويلز (UNSW) في أستراليا، وهي مصممة لغرض تقييم أداء أنظمة اكتشاف الاختراق من خلال توفير بيانات شاملة لحركة مرور الشبكة (traffic) مما يسمح للباحثين بتحليل البيانات على مستويات مختلفة وبالتالي نتائج أكثر دقة لكشف الاختراق [15].

- تحتوي مجموعة البيانات على مجموعة كبيرة من البيانات التي تتضمن أنواعًا مختلفة من أنشطة الشبكة بالإضافة لميزات مثل عناوين IP المصدر والوجهة، ومنافذ المصدر والوجهة، وأنواع البروتوكول، وأحجام الحمولة وغيرها موزعة على 45 عمودًا كما هو موضح في الجدول (2) [16]:

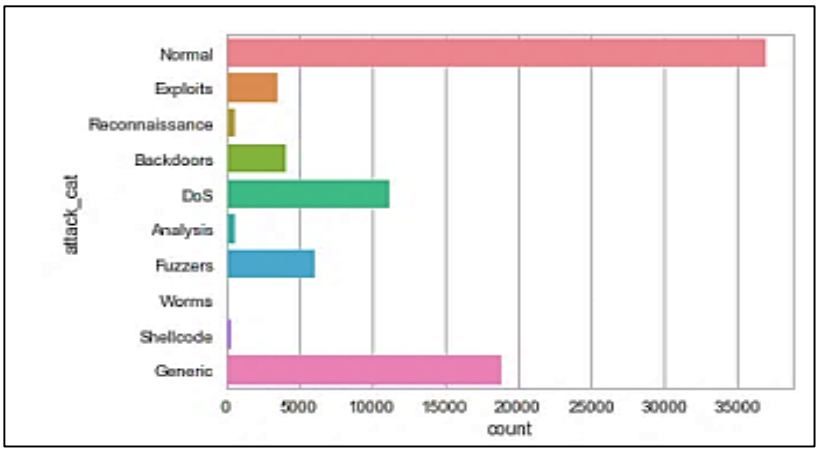
الجدول (2): واصفات مجموعة البيانات UNSW NB15

Attribute Number	Attribute Name	Attribute Number	Attribute Name
1	id	23	dtcpb
2	dur	24	dwin
3	proto	25	tcprtt
4	service	26	synack
5	state	27	ackdat
6	spkts	28	smean
7	dpkts	29	dmean
8	sbytes	30	trans_depth
9	dbytes	31	response_body_len
10	rate	32	ct_srv_src
11	sttl	33	ct_state_ttl
12	dttl	34	ct_dst_ltm
13	sload	35	ct_src_dport_ltm
14	dload	36	ct_dst_sport_ltm
15	sloss	37	ct_dst_src_ltm
16	dloss	38	is_ftp_login
17	sinpkt	39	ct_ftp_cmd
18	dinpkt	40	ct_flw_http_mthd
19	sjit	41	ct_src_ltm
20	djit	42	ct_srv_dst
21	swin	43	is_sm_ips_ports
22	stcpb	44	attack_cat
		45	label

- تنقسم قاعدة البيانات إلى جزء خاص بمرحلة التدريب وجزء خاص بالاختبار وفق التالي:

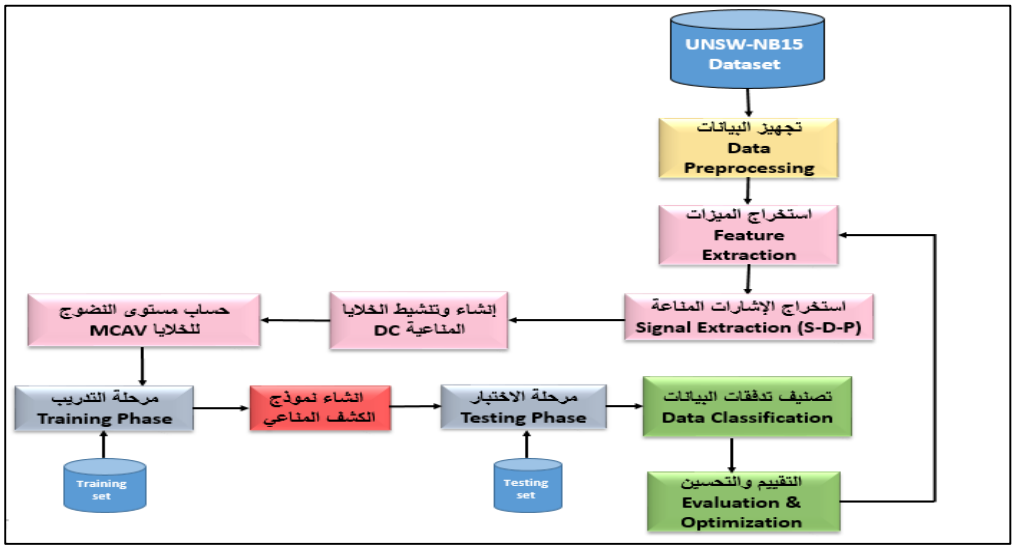
	Anomalous/مشبوهة	Normal/ طبيعية
التدريب (Training): 175341	119.341	65000
الاختبار (Testing): 82332	45332	37000

- كما يضم 9 فئات من الهجمات السيبرانية المتنوعة التالية والموضحة نسبتها في الشكل (4):



٧. آلية العمل:

يوضح الشكل (5) مخطط البحث ومراحل العمل الرئيسية:



الشكل (5): المخطط العام للبحث ومراحل بناء نموذج التصنيف المقترح

٧,١ المعالجة المسبقة للبيانات Data Preprocessing:

بعد تحميل قاعدة البيانات المكونة من 45 واصفة تتنوع قيمها بين عددية ونصية، علينا التحقق من هذه البيانات والتأكد من سلامتها لضمان الحصول على نتائج دقيقة وذلك وفق الخطوات التالية:

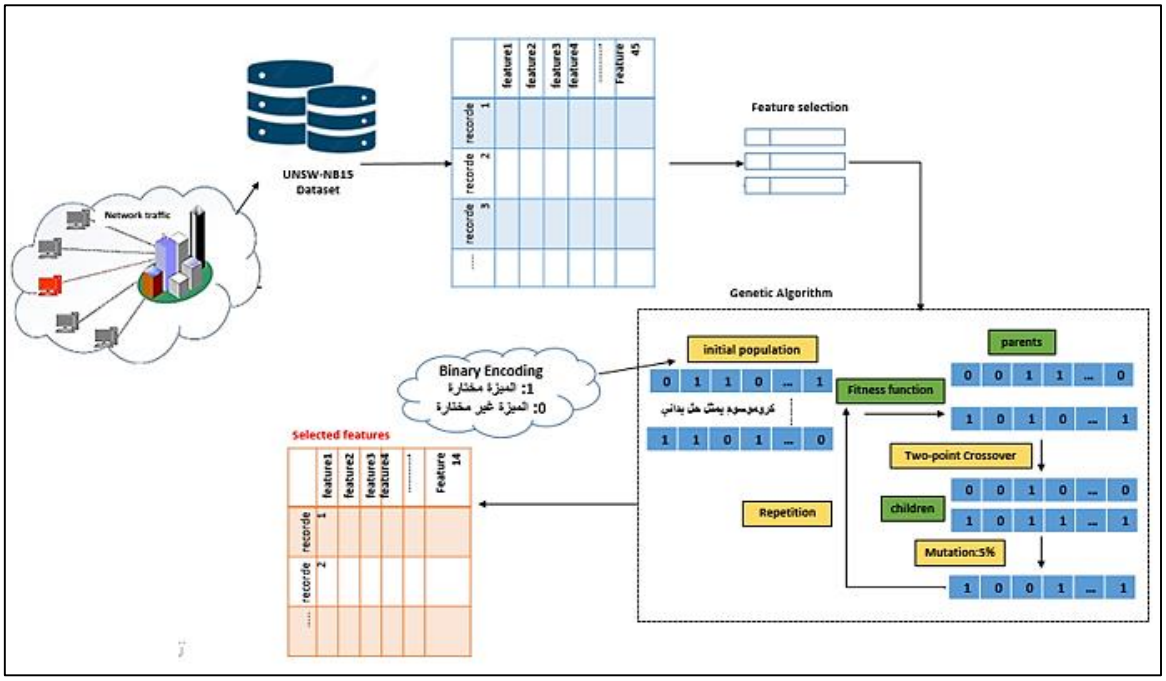
- تم إزالة القيم المفقودة أو تعويضها بالمتوسط الحسابي لقيم الواصفة إن أمكن.

- معالجة القيم المكررة وإزالتها بغرض الحفاظ على تنوع البيانات، حيث لاحظنا وجود بعض السجلات المكررة في واصفة البروتوكول ("proto") وواصفات حركة البيانات ("sttl,dttl").
- تحويل القيم النصية إلى عددية مثل قيم الواصفات ("state, proto, service") لتصبح قابلة للاستخدام من قبل خوارزمية DCA.
- تطبيع البيانات Normalization: حيث هناك بعض الميزات التي تحوي على قيم كبيرة جداً مثل: "sbytes, dbytes, sload, dload" مما يؤثر على أداء الخوارزمية لذلك تم تحويل جميع القيم إلى النطاق [0−1] باستخدام Min−Max Scaling.

٧,٢ استخراج الميزات Feature Extraction:

الغاية من هذه المرحلة هي تحديد الميزات التي تعكس خصائص التدفقات الشبكية وتميز بين الاتصالات العادية والهجمات، وبالتالي منع الواصفات غير المفيدة من التأثير على جودة التصنيف.

قمنا باستخدام الخوارزمية الجينية (Genetic Algorithm) والتي تساعد على اختيار الميزات الأكثر أهمية من خلال محاكاة عملية الانتقاء الطبيعي والعمل على تطوير مجموعة من الحلول المرشحة على مدى أجيال متعددة وذلك بالاعتماد على عملية الاختيار والتكاثر والتطور والطفرة حتى تتقارب الخوارزمية إلى حل مثالي أو شبه مثالي [17]، بحيث نعمل على دمج الخوارزمية البيولوجية هذه مع خوارزمية DCA المناعية بهدف تحسن جودة الحل النهائي وذلك وفق الخطوات الموضحة في الشكل (6) حصلنا بنتيجتها على 13 ميزة رئيسية سنطبق عليها خوارزمية DCA :



الشكل (6): مراحل اختيار الميزات بالاعتماد على الخوارزمية الجينية

بتطبيق الخوارزمية حصلنا على مجموعة الميزات الأكثر ارتباطاً وتأثيراً في مجال الشبكات والموضحة في الجدول التالي:

الجدول (3): واصفات مجموعة البيانات المختارة بعد تطبيق الخوارزمية الجينية

النوع	الميزة	الوصف
واصفات الشبكة Network Attributes	**proto**	نوع البروتوكول (TCP,UDP,ICMP)
	state	حالة الاتصال (SYN_SENT,FIN,ESTABLISH,..)
واصفات حركة البيانات Traffic Attributes	**sbytes**	عدد البايتات المرسلة من المصدر
	dbytes	عدد البايتات المستلمة من الوجهة
	spkts	عدد الحزم المرسلة من المصدر
	dpkts	عدد الحزم المستلمة من الوجهة
واصفات الأداء Performance Attributes	**dur**	مدة الجلسة
	sload	معدل تحميل البيانات من المصدر (bits/s)
	dload	معدل تحميل البيانات إلى الوجهة (bits/s)
واصفات التحكم بالاتصال Connection Control Attribute	**sttl**	قيمة TTL للحزم المرسلة من المصدر
	dttl	قيمة TTL للحزم المستلمة من الوجهة
واصفات المنافذ Ports Attribute	**sport**	رقم منفذ المصدر
	dport	رقم منفذ الوجهة

٧,٣ استخلاص الإشارات المناعيه وتوليد الخلايا الجذعية:

بعد تحديد الميزات الأساسية والأكثر ارتباطاً بعملية كشف التسلل في الشبكة، يتم في هذه المرحلة استخلاص الإشارات المناعية الثلاثة (DS,PAMP,SS) من مجموعة البيانات وتحليلها لإنشاء الخلايا المناعية التي تصنف تدفقات الشبكة، وذلك بالاعتماد على تحليل هذه الميزات السلوكية لحركة البيانات حيث تعبر كل إشارة عن سلوك معين في الشبكة وفق التوزيع في الجدول الآتي:

الجدول (4): توزيع الميزات المختارة على الإشارات ضمن خوارزمية DCA

الإشارة المناعية	الميزات التابعة لها	آلية التأثير
الإشارة الآمنة **Safe Signal (SS)**	**البروتوكول (proto)**	ترتفع قيمة إشارة SS إذا كان البروتوكول tcp مع اتصالات مستقرة.
	حالة الاتصال (state)	ترتفع قيمة SS إذا كانت الحالة CON (اتصال مستمر) أو FIN (إغلاق طبيعي للاتصال).
	عدد الحزم المرسلة والمستقبلة (spkts, dpkts) والبيانات المرسلة (sbytes, dpytes)	كلما كانت قيمها معتدلة وتعكس تفاعل طبيعي ترتفع قيمة SS.
الإشارة الخطرة **Danger Signal (DS)**	**sttl**	كلما كانت قيمتها منخفضة جداً أو متغيرة بشكل غير طبيعي كلما زاد احتمال حدوث هجوم.
	dttl	كلما كانت غير متوافقة مع قيمة sttl لنفس الحزمة كلما زاد احتمال حدوث هجوم أو انتحال الهوية.

	مدة الجلسة (dur)	كلما كانت المدة منخفضة مع عدد حزم مرتفع، كلما زاد احتمال حدوث هجوم.
	منافذ المصدر والوجهة (sport, dport)	كلما كانت القيم مرتفعة أو منخفضة بشكل كبير خارج المجال الطبيعي لها تزداد احتمالية حدوث هجوم وبالأخص DOS.
الإشارة المساعدة (PAMPs)	**معدل تحميل البيانات المرسلة والمستلمة (sload, dload)**	كلما ارتفعت قيمها بشكل غير طبيعي كما أشار ذلك إلى هجوم مؤكد.

بعد حساب قيم الإشارات الثلاث وفق المعادلات (1) و (2) و (3) يتم تمثيل خلايا DC المناعة ضمن الشبكة كمصفوفة مكونة من العناصر الموضحة في الجدول (5) علماً أن قيم الأوزان للإشارات الثلاثة هي كالتالي [18]:

Wight coefficient matrix (WCM) =

Signals	*PAMP*	*SS*	*DS*
CSM	2	1	2
smDC	0	0	2
mDC	2	1	−2

الجدول (5): مكونات مصفوفة الخلية المناعية DC

عنصر خلية DC	**الوصف**
ID: معرّف الخلية	رقم مميز للخلية المناعية
SS: إشارة الأمان	تحدد درجة سلمية النشاط المرتبط بالخلية
DS: إشارة الخطر	تحدد درجة غرابة النشاط المرتبط بالخلية
PAMPs:الإشارة المساعدة	تعبر عن دليل مباشر لوجود تهديد حقيقي
MCAV	مستوى نضوج الخلية
State:حالة الخلية	نشاط عند الاستجابة لنشاط غير طبيعي ، أو خمول في حال عدم تنشيطها، أو شبه ناضجة عند نشاط غير مؤكد

وبتطبيق ما سبق بالاعتماد على لغة بايثون ومكتباتها حصلنا على مجموعة الخلايا المناعة DC التي نوضح عينة منها في الشكل التالي:

```
Cell ID: 1
SS: 0.90, DS: 0.40, PAMP: 0.30
Cell ID: 2
SS: 0.30, DS: 0.70, PAMP: 0.85
Cell ID: 3
SS: 0.50, DS: 0.80, PAMP: 0.75
Cell ID: 4
SS: 0.65, DS: 0.55, PAMP: 0.40
Cell ID: 5
SS: 0.20, DS: 0.90, PAMP: 0.95
```

الشكل (7): توليد الخلايا المناعية DC بناء على قيم الإشارات المناعية المستخلصة

٧,٤ مرحلة التدريب:Training

نقوم في هذه المرحلة بتدريب وتعليم الخلايا المناعة المولدة على تصنيف تدفقات البيانات في الشبكة وتمييز الحركة الطبيعية من الهجومية من خلال تحليل مجموعة البيانات وحساب القيم المناعية لكل ميزة وفق الخطوات الآتية:

١. تحديد قيمة MCAV لكل خليلة مناعية وفق المعادلة (5) بناء على:

- عدد المرات التي تم تصنيف السجل كهجوم (A) N_{attack}.
- إجمالي عدد المرات التي ظهر فيها السجل (A) N_{total} في مجموعة بيانات التدريب.

٢. تحديث الخلايا المناعية DC بناء على قيم MCAV.

٣. تحديد قيم عتبة الشذوذ المناسبة للكشف عن الهجمات، حيث قمنا بالاعتماد على وجود عتبتي شذوذ لزيادة دقة التصنيف وتقليل معدل الإنذارات الكاذبة قدر الإمكان وذلك من خلال المعادلتين المقترحتين (9,10) لحساب قيمة العتبة العليا والدنيا بالاعتماد على قيم المتوسط الحسابي والانحراف المعياري لقيمة MCAV وفق التالي:

$$T_{high} = mean(MCAV) + \alpha * std(mean)$$ **المعادلة[9]**

$$T_{low} = mean(MCAV) - \alpha * std(mean)$$ **المعادلة[10]**

حيث : $mean(MCAV)$: المتوسط الحسابي لقيم MCAV

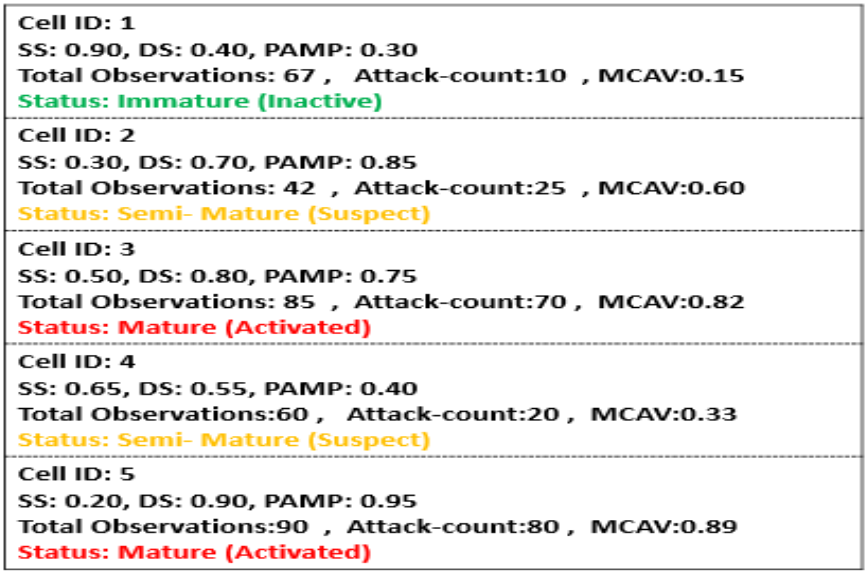
$std(MCAV)$:الانحراف المعياري لقيم MCAV.

α : معامل ضبط Scaling Factor يأخذ قيم ضمن المجال $[1-0]$

٤. تم تحديد حالة الخلايا بالاعتماد على قيم العتبات العليا والدنيا وفق التالي:

قيمة MCAV	حالة الخلية	طبيعة النشاط ضمن الشبكة
[0.0−0.5[	Immune (Inactive)	نشاط طبيعي Benign
[0.5−0.7[	Semi−mature (suspect)	نشاط مشبوه به لكنه غير مؤكد
[0.7−1[	Mature (active)	نشاط عالي الخطورة Malicious

يوضح الشكل (8) نتائج عملية التدريب والقيم الناتجة عنها مع تحديد حالة كل خلية مشكلة:



الشكل (8): نتائج مرحلة التدريب وبناء نموذج التصنيف في خوارزمية DCA

نلاحظ وجود ثلاث حالات للخلية المناعية مما ينتج عنها ثلاث نتائج للتصنيف وهذا يتعارض مع مجموعة البيانات UNSW التي تتضمن تصنيفين فقط وهما 0: Benign و 1: Malicious لذلك علينا معالجة الصنف الثالث المشبوه (suspect) ليندرج ضمن أحد التصنيفين السابقين.

لمعالجة هذا الأمر لجئنا إلى قيمة إشارة الخطر DS لحسم النتيجة كونها تعكس مدى خطورة النشاط الشبكي ومؤشر قوي على الهجمات الحقيقية مما يسمح بالتقليل من الإنذارات الكاذبة False Positive وزيادة دقة التصيف وفق التالي:

- إذا كانت قيمة DS للخلية المناعية أكبر أو يساوي 0.7 عندئذ تُصنف على أنّها خبيثة Malicious.
- إذا كانت قيمة DS للخلية المناعية أصغر من 0.7 عندئذ تُصنف على أنّها سليمة Benign.

٥. حفظ نموذج الكشف المكون من مجتمع الخلايا المناعية لاستخدامه في عملية الاختبار وذلك من خلال تخزين كافة قيم البيانات الناتجة في هذه المرحلة.

٧,٥ مرحلة الاختبار Testing:

بعد تدريب الخوارزمية باستخدام مجموعة بيانات التدريب وبناء نموذج الكشف المناسب، الآن سنقوم باختبار نموذج الكشف الناتج على مجموعة بيانات الاختبار المكونة من 82,332 سجل وتقييم النتائج بالاعتماد على مجموعة معايير التقييم المذكورة سابقاً وفق التالي:

١- تحميل نموذج الكشف المحفوظ سابقاً والذي يحوي على الخلايا المناعية المدربة وقيم العتبات العليا والدنيا.

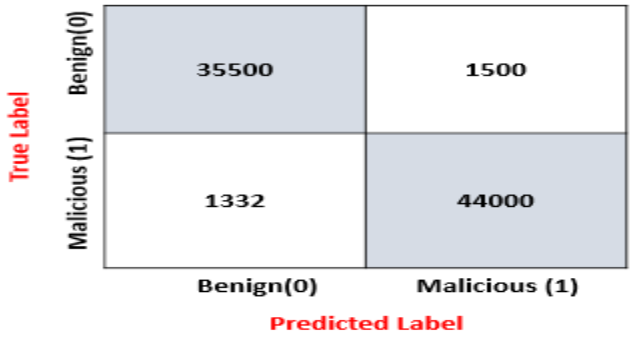
٢- تمرير مجموعة بيانات الاختبار إلى مجتمع الخلايا المناعية لحساب قيم الإشارات المناعية SS,DS,PAMP لكل سجل جديد.

٣- حساب قيمة MCAV لكل سجل باستخدام قيم الإشارات الثلاث.

٤- مقارنة قيم MCAV الناتجة مع العتبات لتصيف السجل إلى إحدى حالتي التصنيف (طبيعي، شاذ)

٧,٥,١ تقييم نتائج الاختبار:

بتطبيق نموذج الكشف المقترح على مجموعة بيانات الاختبار وفق ما ذكرنا سابقاً من خطوات حصلنا على مصفوفة الارتباك Conclusion Matrix الموضحة في الشكل (9):



الشكل (9): مصفوفة الارتباك الناتجة عن تطبيق نموذج التصنيف المقترح على مجموعة بيانات الاختبار

بالاعتماد على قيم المصفوفة وهي:

TP=44000, FP=1500, FN=1332, TN=35500

- قمنا بحساب قيم معايير التصنيف وفق المعادلات (5,6,7,8) لنحصل على

النتائج الموضحة في الجدول (6):

الجدول (6): قيم معايير التقييم الناتجة عن تطبيق نموذج الكشف المقترح على مجموعة بيانات الاختبار

Evaluation Matrices Values						
Accuracy	precision	Recall	F1-Score	RMSE	DR	FAR
97%	96.7%	97.06%	96.88%	0.158	97,27%	4%

كما هو معروف فإنّ من أهم العوامل التي تؤثر على أداء نموذج الكشف والتي تعتبر مقياس رئيسي في تحديد مدى فعاليته على الأنظمة الأمنية هو عامل الزمن الذي يستغرقه النموذج في كشف الهجمات وتنبيه مختصي الأمن لأن هذه تطبيقات Real-time ويجب أن تعمل في الوقت الفعلي وتنجز المطلوب منها بسرعة وزمن صغير جداً لتكون قادرة على أداء المهمة المطلوبة منها بجودة عالية وإلا مهما بلغت دقة تصنيفه يعتبر فاشل.

أهم العوامل التي تؤثر على زمن تنفيذ النموذج المقترح وتدخل في حسابه (بغض النظر عن طبيعة الموارد ومواصفات الأجهزة المستخدمة والتي أيضاً لها دور فعال في تحسين الزمن) هي كالتالي:

العامل المؤثر	الزمن التقديري (s)
١. استخراج الميزات	1.28
٢. توليد الخلايا المناعية	1.83
٣. حساب الإشارات المناعية	1.6
٤. التصنيف	2.29
إجمالي زمن التصنيف	**7 sec**

٨. النتائـــج والمناقشـــة:

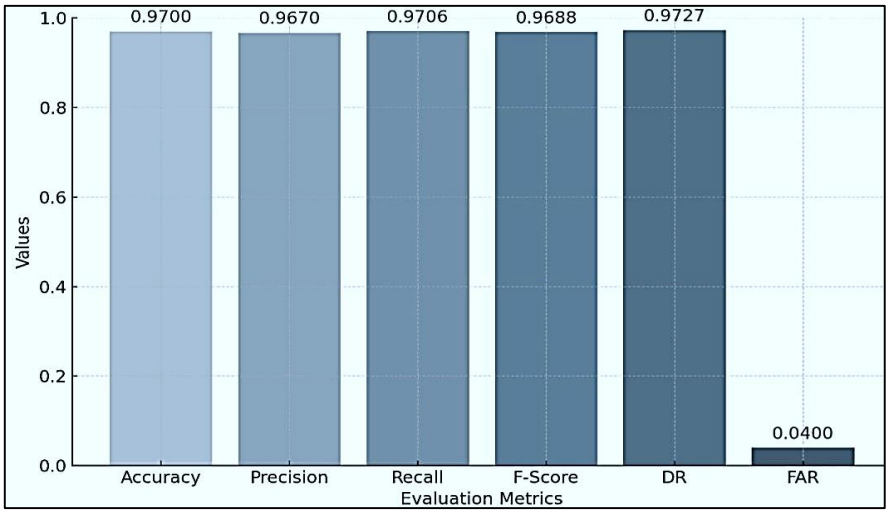
توضح النتائج النهائية لتصنيف مجموعة بيانات الاختبار وكشف الهجمات في الشبكة والتي توصلنا إليها بالاعتماد على خوارزمية الخلايا الجذعية المناعية DCA والخوارزمية الجينية GA أنّ نموذج التصنيف المقترح حقق معدل دقة مرتفع في تصنيف الهجمات وصل إلى 97% مما يجعله نظام مناسب لتأمين الشبكات الحاسوبية بالإضافة إلى نسبة مرتفعة لقيمة DR والذي يمثل نسبة الهجمات المكتشفة بشكل صحيح حيث وصل إلى 97.27% مما يؤكد على قدرة النموذج في التعامل مع معظم الهجمات، بينما هناك عدد صغير من الإنذارات الخاطئة (كما نلاحظ في الشكل (10)) من قبله ناتجة عن تصنيف الحركة الطبيعية للبيانات كهجوم وتعتبر قيمتها مقبولة في معظم الأنظمة الأمنية .

تعطي قيمة RMSE المنخفضة مؤشر حقيقي على مدى موثوقية النموذج المقترح مع قيمة وصلت إلى 0.158 مما يضمن دقة التصنيف وموثوقيته في بيئة الشبكة الحقيقية.

نلاحظ وجود توازن مثالي بين تقليل الإنذارات الكاذبة (FP) وعدم فقدان هجمات حقيقية (FN) ، مع ضمان الحفاظ على التوازن بين الدقة والحساسية كما يشير معيار التقييم F1-Score والذي وصل إلى 96.88% ، مما يعكس الأداء الكفاءة العالية للنموذج المقترح .

بمراقبة عامل الزمن نجد أنّ عملية التصنيف لكافة السجلات في مجموعة الاختبار قد تمت بزمن منخفض وصل إلى 0.085ms لكل سجل مما يجعل النموذج المقترح مناسب للعمل في بيئات الزمن الحقيقي real-time .

عالج النظام المقترح مشكلة تصنيف حركة البيانات غير الاعتيادية إلى هجوم مباشرة دون تحليلها بشكل دقيق وهذا ما كان يزيد مع معدل الإنذارات الكاذبة في التقنيات التقليدية، وذلك من خلال الاعتماد على فئة النشاط المشتبه به Suspect وتصنيفه وفق درجة الخطورة مع الأخذ بعين الاعتبار قيم إشارة PAMP الداعمة والمساعدة في اتخاذ القرار النهائي.



الشكل (10): قيم معايير التقييم الناتجة عن تطبيق نموذج الكشف المقترح على مجموعة بيانات

٩. الاســـتنتاجات والتوصيـــات:

مع التطور التكنولوجي الكبير واعتماد الأشخاص والمؤسسات والحكومات على الانترنت للقيام بأعمالهم وتقديم خدماتهم أصبحت تشكل عملية تأمين الأنظمة الشبكية تحدياً كبيرا يتطلب العمل بشكل جدي على استثمار هذه التكنولوجيا ومخرجاتها في بناء منظومة أمنية تعمل على حماية الشبكات والبيانات المتنقلة خلالها والخوادم والأجهزة المتصلة بها توفر دقة عالية وموثوقية كبيرة.

قدمنا خلال هذا البحث باقتراح نموذج هجين لكشف وتصنيف الهجمات في الشبكات الحاسوبية بالاعتماد على الأنظمة البيولوجية الذكية وبالأخص خوارزمية الخلايا الجذعية المناعية (DCA)والخوارزمية الجينية التي تعتبر من الأنظمة الذكية الواعدة في مجال الهندسة الحيوية وذلك بالاعتماد على مجموعة البيانات–UNSW NB15المعيارية في مرحلة بناء النموذج وتدريبه واختباره.

وبدراسة أداء النموذج المناعي المقترح في كشف هجمات الشبكة وتحليل حركة البيانات ضمنها أظهرت النتائج قدرته العالية على تصنيف الهجمات والتعرف عليها بدقة وكفاءة مع ضمان معدل منخفض للإنذارات الكاذبة.

هناك مجموعة من التحديات التي تعترض عملية تطبيق هذه الخوارزميات الحيوية في مجال الأمن السيبراني وهو العدد الكبير للمعاملات الرياضية التي تؤثر على أداء الخوارزمية مما يتطلب ضبط دقيق لقيمها ضمن مجال محدد يضمن التأثير الإيجابي على الخوارزمية مثل عتبة التصنيف العليا T_{high}والدنيا T_{low} وقيمة الفا α، بالإضافة إلى معاملات الخوارزمية الجينية المستخدمة في استخلاص الميزات التي تؤثر بشكل رئيسي على نتيجة التصنيف مثل عدد الأجيال ونسبة الطفرة وطرق التصالب والانتقاء. هذا كله يحتاج لتجارب عديدة تشمل تطبيق قيم مختلفة لكل من المعاملات السابقة لنتمكن من تحديد أي القيم الأكثر فعالية ومدى تأثير كل منها على عملية التصنيف النهائية وبالتالي دقة النموذج المقترح.

وبالنتيجة يقترح البحث التوصيات التالية:

- ❖ الاعتماد على خوارزميات التعلم الآلي في التنبؤ بقيم المعاملات المناسبة مما يؤثر إيجاباً على أداء النموذج ودقة نتائج التصنيف.
- ❖ اختبار نموذج التصنيف مع حجم مجموعة بيانات أكبر ودراسة مدى فعاليتها مع زيادة حجم البيانات.
- ❖ دراسة تأثير القيم المختلفة لمعاملات الخوارزمية على قيم FP و FN.
- ❖ دراسة تأثير ارتفاع او وانخفاض قيمة العتبة المستخدمة في فرز الصنف الثالث Suspect للبيانات وتأثير ذلك على كفاءة التصنيف.

١٠. المراجع:

[1] Li, L., Lin, Q., & Ming, Z. (2022). A survey of artificial immune algorithms for multi-objective optimization. Neurocomputing, 489, 211-229

[2] Dasgupta, D., Yu, S., & Nino, F. (2011). Recent advances in artificial immune systems: models and applications. Applied Soft Computing, 11(2), 1574-1587.

[3] Datta Gupta, K., & Dasgupta, D. (2021). Negative Selection Algorithm Research and Applications in the last decade: A Review. arXiv e-prints, arXiv-2105.

[4] Othman, S. M., Alsohybe, N. T., Ba-Alwi, F. M., & Zahary, A. T. (2018). Survey on intrusion detection system types. International Journal of Cyber-Security and Digital Forensics, 7(4), 444-463.

[5] Ashiku, L., & Dagli, C. (2021). Network intrusion detection system using deep learning. Procedia Computer Science, 185, 239-247.

[6] Chelly, Z., & Elouedi, Z. (2016). A survey of the dendritic cell algorithm. Knowledge and Information Systems, 48, 505-535.

[7] Greensmith, J. (2007). The dendritic cell algorithm (Doctoral dissertation, University of Nottingham).

[8] Pereira, V. E. (2024). Automatic signal characterization for the Dendritic Cell Algorithm.

[9] Pereira, G. (2011). Artificial Immune System Algorithm based on Danger Theory.

[10] Dagdia, Z. C. (2019). A scalable and distributed dendritic cell algorithm for big data classification. Swarm and Evolutionary Computation, 50, 100432.

[11] Dagdia, Z. C. (2019). A scalable and distributed dendritic cell algorithm for big data classification. Swarm and Evolutionary Computation, 50, 100432.

[12] Muntean, M., & Militaru, F. D. (2023, January). Metrics for evaluating classification algorithms. In Education, Research and Business Technologies: Proceedings of 21st International Conference on Informatics in Economy (IE 2022) (pp. 307-317). Singapore: Springer Nature Singapore.

[13] Vujović, Ž. (2021). Classification model evaluation metrics. International Journal of Advanced Computer Science and Applications, 12(6), 599-606.

[14] Salih, A. A., & Abdulazeez, A. M. (2021). Evaluation of classification algorithms for intrusion detection system: A review. Journal of Soft Computing and Data Mining, 2(1), 31-40.

[15] Kanimozhi, V., & Jacob, P. (2019). UNSW-NB15 dataset feature selection and network intrusion detection using deep learning. International Journal of Recent Technology and Engineering, 7(5), 443-446.

[16] Moustafa, N., & Slay, J. (2016). The evaluation of Network Anomaly Detection Systems: Statistical analysis of the UNSW-NB15 data set and the comparison with the KDD99 data set. Information Security Journal: A Global Perspective, 25(1-3), 18-31.

[17] Sohail, A. (2023). Genetic algorithms in the fields of artificial intelligence and data sciences. Annals of Data Science, 10(4), 1007-1018.

[18] Greensmith (2007) The dendritic cell algorithm (Ph.D. thesis edn). Citeseer.