Comprehensive Journal of Science
Volume (11), Issue (42), (Sapt, 2026)
ISSN: 3104-6266

مجلة العلوم الشاملة
المجلد (11) العدد 42 سبتمبر 2026
ردمد: 3104-6266

تقييم استقرار وكفاءة اختيار الخصائص في أنظمة كشف التسلل متعددة الفئات باستخدام خوارزميات التعلم الآلي

الهام يخلف أبوالشواشي
قسم التقنية الالكترونية ـ كلية التقنية الهندسية زوارة ـ زوارة ليبيا

Evaluating Feature Selection Stability and Efficiency in Multiclass Intrusion Detection Systems Using Machine Learning Algorithms

Elham Yekhlef Abushwashi
Department of Electronic Technology - Faculty of Engineering Technology, Zuwarah – Libya
elham_abuelshwashi@yahoo.com

تاريخ الاستلام: 2026/09/9 –تاريخ المراجعة: 2026/09/10– تاريخ القبول: 2026/10/8– تاريخ للنشر: 6/11/ 2026

الملخص

تهدف هذه الدراسة إلى تقييم استقرار وكفاءة خوارزميات اختيار الخصائص في أنظمة كشف التسلل متعددة الفئات باستخدام خوارزميات التعلم الآلي، وذلك من خلال دراسة أثر اختيار الخصائص على استقرار المجموعات المختارة وأداء التصنيف والكفاءة الحسابية. استخدمت الدراسة مجموعة بياناتUNSW-NB15 ، وتم تطبيق خوارزميتي المعلومات المتبادلة (Mutual Information – MI)واختبار كاي تربيع (Chi-Square) لاختيار أفضل 20 خاصية من أصل 194 خاصية. وتم تقييم استقرار الاختيار باستخدام معامل جاكارد عبر خمس عينات عشوائية، كما تم تقييم أداء التصنيف باستخدام مقاييس Accuracyو Precisionو RecallوF1-score، إلى جانب قياس زمن التدريب والتنبؤ.

أظهرت النتائج أن خوارزمية MI حققت أعلى استقرار في اختيار الخصائص بمتوسط معامل جاكارد بلغ 1.0000، مقارنة بـ0.8883 لخوارزمية Chi-Square. كما حققت MI أعلى قيمة لمقياس Macro F1-score بلغت 0.5226، وأعلى دقة كلية بلغت 0.7487. في المقابل، حققت خوارزمية Chi-Square أفضل كفاءة زمنية، حيث بلغ زمن التدريب 326.42 ثانية وزمن التنبؤ 1.09 ثانية. وتشير النتائج إلى أهمية تقييم خوارزميات اختيار الخصائص من جوانب متعددة تشمل الاستقرار وجودة التصنيف والكفاءة الحسابية، وعدم الاعتماد على مقياس واحد لتحديد الخوارزمية الأنسب لأنظمة كشف التسلل متعددة الفئات.

الكلمات المفتاحية :أنظمة كشف التسلل، اختيار الخصائص، التعلم الآلي، التصنيف متعدد الفئات، معامل جاكارد، UNSW-NB15.

Abstract
This study aims to evaluate the stability and efficiency of feature selection algorithms in multi-class intrusion detection systems (IDS) that utilize machine learning. It examines the impact of feature selection on the stability of the selected feature subsets, classification performance, and computational efficiency. Using the UNSW-NB15 dataset, the study applied the Mutual Information (MI) and Chi-Square algorithms to select the top 20 features out of 194. Selection stability was assessed using the Jaccard coefficient across five random samples, while classification performance was evaluated using metrics such as Accuracy, Precision, Recall, and F1-score. Additionally, training and prediction times were measured.
The results indicate that the MI algorithm achieved the highest stability in feature selection, with an average Jaccard coefficient of 1.0000, compared to 0.8883 for the Chi-Square algorithm. MI also yielded the highest Macro F1-score (0.5226) and the highest overall accuracy (0.7487). Conversely, the Chi-Square algorithm demonstrated superior temporal efficiency, with a training time of 326.42 seconds and a prediction time of 1.09 seconds. These findings underscore the importance of evaluating feature selection algorithms across multiple dimensions—including

stability, classification quality, and computational efficiency—rather than relying on a single metric to determine the most suitable algorithm for multi-class intrusion detection systems. Keywords: Intrusion detection systems, feature selection, machine learning, multi-class classification, Jaccard coefficient, UNSW-NB15.

1. المقدمة

أصبحت شبكات الحاسوب عرضة بصورة متزايدة لأنماط متعددة من الهجمات السيبرانية، مما جعل أنظمة كشف التسلل (Intrusion Detection Systems – IDS)أحد المكونات الأساسية في أمن الشبكات. ومع ازدياد حجم بيانات حركة الشبكة وتنوع خصائصها، أصبح الاعتماد على جميع الخصائص المتاحة في عملية التصنيف يؤدي إلى زيادة العبء الحسابي، وقد يتضمن خصائص غير مؤثرة أو متكررة.

توفر تقنيات التعلم الآلي إمكانية تحليل حركة الشبكة والتعرف على الأنماط المرتبطة بالسلوك الطبيعي والهجمات. وفي المقابل، يمكن لتقنيات اختيار الخصائص (Feature Selection) تقليل عدد الخصائص المستخدمة مع المحافظة على المعلومات الأكثر أهمية للكشف . وقد أظهرت الدراسات الحديثة استمرار الاهتمام بدمج اختيار الخصائص مع التعلم الآلي في أنظمة كشف التسلل، خصوصًا لمعالجة ارتفاع أبعاد بيانات الشبكة وعدم توازن الفئات

2. مشكلة البحث

تُعد أنظمة كشف التسلل القائمة على التعلم الآلي من الأساليب المهمة في اكتشاف الأنشطة والهجمات غير الطبيعية في شبكات الحاسوب، إلا أن فعالية هذه الأنظمة تتأثر بصورة كبيرة بخصائص البيانات المستخدمة في بناء نماذج الكشف. وتحتوي مجموعات بيانات حركة الشبكات الحديثة، مثلUNSW-NB15 ، على عدد كبير من الخصائص التي تصف جوانب مختلفة من الاتصالات الشبكية، الأمر الذي قد يؤدي إلى زيادة التعقيد الحسابي، وإدخال خصائص غير مهمة أو متكررة، والتأثير في قدرة نماذج التعلم الآلي على التمييز بين الفئات المختلفة للهجمات.

وعلى الرغم من استخدام العديد من أساليب اختيار الخصائص بهدف تقليل أبعاد البيانات وتحسين أداء أنظمة كشف التسلل، فإن اختيار مجموعة الخصائص الأكثر ملاءمة لا يعتمد فقط على قدرتها على تحسين دقة التصنيف؛ إذ قد تختلف الخصائص المختارة باختلاف طريقة الاختيار المستخدمة، كما أن مجموعة الخصائص التي تحقق أداءً جيدًا مع نموذج تصنيف معين قد لا تحقق الأداء نفسه مع نماذج أخرى. وتظهر الدراسات الحديثة اهتمامًا متزايدًا بتقليل عدد الخصائص وتحسين الكفاءة، إلا أن المقارنة المنهجية بين استقرار الخصائص المختارة وتأثير ذلك في الأداء متعدد الفئات والكلفة الحسابية ما تزال تحتاج إلى مزيد من الدراسة.

وعليه، تتمثل مشكلة البحث في عدم وضوح مدى استقرار الخصائص التي يتم اختيارها بواسطة أساليب مختلفة لاختيار الخصائص، ومدى قدرة مجموعة خصائص مستقرة ومختصرة على المحافظة على كفاءة نظام كشف التسلل متعدد الفئات مع تقليل عدد الخصائص والتكلفة الحسابية للنماذج.

ومن ثم يسعى البحث إلى دراسة العلاقة بين استقرار اختيار الخصائص، وعدد الخصائص المختارة، وأداء نماذج التعلم الآلي، والكفاءة الحسابية، وذلك من خلال تطبيق ومقارنة عدة أساليب لاختيار الخصائص على مجموعة بيانات-UNSW NB15، ثم اختبار الخصائص المختارة باستخدام مجموعة من خوارزميات التصنيف متعددة الفئات

3. اسئلة البحث

يسعى البحث إلى الإجابة عن الأسئلة الآتية:

1. ما مدى التوافق والاستقرار بين الخصائص التي يتم اختيارها باستخدام أساليب اختيار الخصائص المختلفة من بيانات UNSW-NB15؟
2. إلى أي مدى يؤثر تقليل عدد الخصائص على أداء نماذج التعلم الآلي في كشف وتصنيف هجمات الشبكات متعددة الفئات؟
3. هل يؤدي استخدام مجموعة خصائص مختارة إلى تحسين الكفاءة الحسابية لنظام كشف التسلل، من حيث زمن التدريب والتنبؤ، مع المحافظة على مستوى مناسب من أداء الكشف؟
4. ما مجموعة الخصائص وطريقة اختيار الخصائص التي تحقق أفضل توازن بين استقرار الخصائص، وأداء التصنيف، والتعقيد الحسابي؟

4. أهداف البحث

يهدف هذا البحث إلى تحقيق الأهداف الآتية:

1. تقييم ومقارنة أساليب اختيار الخصائص لتحديد الخصائص الأكثر أهمية واستقرارًا في بيانات كشف التسلل متعددة الفئات.

2. دراسة تأثير اختيار الخصائص وتقليل عددها على أداء نماذج التعلم الآلي في تصنيف أنواع الهجمات المختلفة.
3. تقييم كفاءة نماذج كشف التسلل من خلال مقارنة الأداء والدقة ومعدلات الخطأ، إلى جانب زمن التدريب والتنبؤ قبل وبعد اختيار الخصائص.
4. تحديد مجموعة خصائص مختصرة ومستقرة تحقق توازنًا مناسبًا بين أداء الكشف والتعقيد الحسابي، ويمكن الاعتماد عليها في بناء نظام كشف تسلل متعدد الفئات.

5. أهمية البحث

تنبع أهمية هذا البحث من الحاجة إلى تطوير أنظمة كشف تسلل أكثر كفاءة وقدرة على التعامل مع تعدد أنواع الهجمات، مع الحد من التعقيد الناتج عن استخدام عدد كبير من خصائص حركة الشبكات. وتتمثل أهمية البحث في الجوانب الآتية:

1. **الأهمية العلمية:**
يساهم البحث في دراسة استقرار أساليب اختيار الخصائص ومقارنة مدى توافق الخصائص المختارة بينها، بما يوفر فهمًا أفضل لدور اختيار الخصائص في أنظمة كشف التسلل متعددة الفئات.
2. **الأهمية التطبيقية:**
يساعد تحديد مجموعة خصائص مختصرة ومستقرة في بناء نماذج كشف تسلل تتطلب موارد حسابية أقل، مع المحافظة على مستوى مناسب من أداء التصنيف.
3. **أهمية تحسين الأداء:**
يوفر البحث مقارنة منهجية بين تأثير الخصائص الأصلية والخصائص المختارة على أداء مجموعة من خوارزميات التعلم الآلي، بما يساعد في تحديد الاختيارات الأكثر ملاءمة لبناء أنظمة كشف التسلل.
4. **الأهمية المستقبلية:**
يمكن أن تشكل النتائج أساسًا لدراسات لاحقة تتناول تطوير أنظمة كشف تسلل أكثر خفة وكفاءة، أو اختبار استقرار الخصائص المختارة على مجموعات بيانات وبيئات شبكية مختلفة.

6. الدراسات السابقة

تناولت دراسة More et al. (2024) تحسين أداء أنظمة كشف التسلل باستخدام مجموعة بيانات UNSW-NB15 ، من خلال تطبيق خوارزميات Logistic Regression و SVM و Decision Tree و Random Forest، مع استخدام تحليل الارتباط والعينات العشوائية لاختيار الخصائص. وأظهرت النتائج تفوق Random Forest من حيث دقة التصنيف ومقياس F1، مما يبرز أهمية اختيار الخصائص في تحسين أداء أنظمة كشف التسلل .

واقترحت دراسة Sowmya and Mary Anita (2025) خوارزمية IV-RFE لاختيار الخصائص، بالاعتماد على التباين النسبي وأوزان الخصائص و Recursive Feature Elimination، مع التركيز على تحقيق الاستقرار في الخصائص المختارة وتحسين دقة الكشف عن الهجمات. وأكدت الدراسة أهمية الاستقرار بوصفه معيارًا إضافيًا عند تقييم أساليب اختيار الخصائص .

وقدمت دراسة Umer et al. (2025) نموذجًا لكشف التسلل باستخدام مجموعة بيانات UNSW-NB15 ، يجمع بين اختيار الخصائص المعتمد على التغليف (Wrapper-Based Feature Selection) ونموذج Multi-Head Attention Transformer. وهدفت الدراسة إلى تقليل عدد الخصائص وتحسين أداء التصنيف، وحققت نتائج بلغت 93% للدقة الكلية و92% لمقياس F1.

أما دراسة Saputra et al. (2026) فقد اقترحت أسلوبًا هجينًا لاختيار الخصائص يسمى IGL1 ، يجمع بين Information Gain و L1-Regularization، واختبرته باستخدام عدة خوارزميات تعلم آلي على مجموعة بيانات UNSW-NB15 بعد تقليصها إلى ست فئات. وحقق نموذج Random Forest أفضل النتائج بدقة بلغت 84.32% ومقياس F1 بلغ 83.76 %.

يتضح من الدراسات السابقة أن اختيار الخصائص يمثل أحد الأساليب المهمة لتحسين أداء أنظمة كشف التسلل وتقليل أبعاد بيانات حركة الشبكات، كما أظهرت الدراسات الحديثة اهتمامًا متزايدًا بتقليل عدد الخصائص وتحسين الكفاءة الحسابية. كذلك برز في بعض الدراسات مفهوم استقرار الخصائص باعتباره عاملًا مهمًا إلى جانب أداء التصنيف. ومع ذلك، يختلف تركيز الدراسات السابقة بين تحسين أداء نموذج تصنيف محدد، أو تطوير طريقة جديدة لاختيار الخصائص، أو تقليل عدد الخصائص، مع محدودية المقارنة المنهجية بين استقرار الخصائص الناتجة عن أساليب اختيار مختلفة.

7. الفجوة البحثية

ظهرت الدراسات السابقة اهتمامًا متزايدًا باستخدام أساليب اختيار الخصائص في أنظمة كشف التسلل، بهدف تقليل أبعاد البيانات وتحسين أداء نماذج التعلم الآلي وخفض التكلفة الحسابية. ومع ذلك، لا يزال تقييم استقرار الخصائص المختارة، ومدى توافق نتائج أساليب الاختيار المختلفة، وعلاقة ذلك بأداء التصنيف والكفاءة الحسابية، بحاجة إلى مزيد من الدراسة.

كما ركزت بعض الدراسات على تحسين دقة التصنيف أو تقييم أداء نماذج محددة، دون تناول العلاقة المتكاملة بين عدد الخصائص المختارة، واستقرارها، وأداء التصنيف متعدد الفئات، وزمن التدريب والتنبؤ.

وعليه، تتمثل الفجوة البحثية في الحاجة إلى تقييم مقارن لأساليب اختيار الخصائص، يجمع بين قياس استقرار الخصائص المختارة، وتقييم تأثيرها على أداء نماذج التعلم الآلي وكفاءتها الحسابية في تصنيف أنواع متعددة من الهجمات. ويسعى هذا البحث إلى معالجة هذه الفجوة من خلال مقارنة أسلوبي المعلومات المتبادلة (MI) ومربع كاي(Chi-Square) ، بهدف تحديد مدى قدرتهما على اختيار مجموعة مختصرة ومستقرة من الخصائص تحقق توازنًا مناسبًا بين جودة التصنيف وتقليل التعقيد الحسابي.

8. أنظمة كشف التسلل وأهميتها في الأمن السيبراني

تُعد أنظمة كشف التسلل (Intrusion Detection Systems – IDS) من المكونات الأساسية في أمن الشبكات، إذ تهدف إلى مراقبة الأنشطة داخل الأنظمة الحاسوبية وحركة البيانات عبر الشبكات، وتحليلها للكشف عن محاولات الوصول غير المصرح به أو الأنشطة التي قد تمثل تهديدًا لأمن المعلومات. وتساعد هذه الأنظمة في اكتشاف السلوكيات المشبوهة وإصدار التنبيهات اللازمة، بما يدعم اتخاذ الإجراءات المناسبة للحد من تأثير الهجمات الإلكترونية وحماية سرية المعلومات وسلامتها وتوافرها.

وتزداد أهمية هذه الأنظمة مع التطور المستمر في أساليب الهجمات الإلكترونية وتنوعها، مما يتطلب تطوير آليات أكثر قدرة على تحليل حركة الشبكة والتمييز بين الأنشطة الطبيعية والضارة، مع مراعاة دقة الكشف وسرعته وتقليل الإنذارات الخاطئة. ويمكن تصنيف أنظمة كشف التسلل وفقًا لآلية الكشف إلى نوعين رئيسيين:

أولًا: أنظمة كشف التسلل المعتمدة على التوقيعات:

تعتمد على مقارنة الأنشطة المرصودة بتوقيعات وأنماط معروفة مسبقًا للهجمات، وتتميز بقدرتها على اكتشاف الهجمات المعروفة، إلا أن قدرتها على اكتشاف الهجمات الجديدة تظل محدودة ما لم تُحدَّث قاعدة التوقيعات باستمرار.

ثانيًا: أنظمة كشف التسلل المعتمدة على الشذوذ:

تعتمد على تحديد أنماط السلوك الطبيعي للشبكة، ثم اكتشاف الأنشطة التي تنحرف عنها. وتتيح هذه الآلية إمكانية التعرف على بعض الهجمات غير المعروفة مسبقًا، لكنها قد تواجه صعوبة في التمييز بين الأنشطة غير المعتادة والأنشطة الضارة، مما قد يؤدي إلى زيادة الإنذارات الخاطئة.

وقد أسهم تطور تقنيات التعلم الآلي في تعزيز إمكانات الكشف المعتمد على الشذوذ، من خلال تحليل بيانات حركة الشبكة واستخلاص الأنماط التي تساعد على التمييز بين الأنشطة الطبيعية والهجومية.

9. أنظمة كشف التسلل متعددة الفئات

يشير التصنيف متعدد الفئات إلى قدرة نظام كشف التسلل على تصنيف حركة الشبكة إلى عدة فئات محددة، بدلًا من الاقتصار على التمييز بين الحركة الطبيعية والحركة الهجومية. ففي التصنيف الثنائي تُجمع الهجمات المختلفة ضمن فئة واحدة، بينما يتيح التصنيف متعدد الفئات تحديد نوع الهجوم المحتمل.

وتكمن أهمية هذا النوع من التصنيف في اختلاف الهجمات الإلكترونية من حيث خصائصها وسلوكها وتأثيرها في الشبكة. كما يتيح تقييم قدرة نماذج التعلم الآلي على التمييز بين أنواع متعددة من الهجمات، ويوفر معلومات أكثر تفصيلًا عن طبيعة التهديدات الأمنية.

10. التعلم الآلي في أنظمة كشف التسلل

يُعد التعلم الآلي من التقنيات المستخدمة في تطوير أنظمة كشف التسلل، إذ يتيح بناء نماذج قادرة على استخلاص الأنماط والعلاقات من بيانات حركة الشبكة، والاستفادة منها في التمييز بين الأنشطة الطبيعية والأنشطة المرتبطة بالهجمات الإلكترونية.

وتعتمد فعالية هذه النماذج على جودة البيانات المستخدمة في التدريب، ومدى ملاءمة الخصائص المدخلة، وقدرة الخوارزمية على التمييز بين الفئات المختلفة. كما تتأثر كفاءتها بحجم البيانات وعدد الخصائص المستخدمة، الأمر الذي يجعل اختيار الخصائص المناسبة خطوة مهمة لتحسين عملية التعلم وتقليل المتطلبات الحسابية.

ومن أبرز خوارزميات التعلم الآلي المستخدمة في تصنيف بيانات الشبكات:

- **شجرة القرار :(Decision Tree)** تعتمد على بناء قواعد متفرعة تُستخدم في تصنيف البيانات، وتمتاز بسهولة تفسير نتائجها نسبيًا.
- **الغابة العشوائية :(Random Forest)** تعتمد على دمج نتائج مجموعة من أشجار القرار بهدف تحسين أداء التصنيف وتقليل الاعتماد على نموذج منفرد.
- **آلة المتجهات الداعمة :(Support Vector Machine – SVM)** تعتمد على تحديد حدود فاصلة بين الفئات، ويمكن استخدامها في مسائل التصنيف المختلفة.

- **التعزيز المتدرج :(XGBoost)** يعتمد على بناء نماذج متتابعة تعمل على تحسين الأداء تدريجيًا من خلال معالجة أخطاء النماذج السابقة.

وتتأثر فعالية هذه الخوارزميات بمدى ملاءمة الخصائص المستخدمة لتمثيل حركة الشبكة، إذ قد يؤدي استخدام خصائص غير ضرورية أو متكررة إلى زيادة العبء الحسابي دون تحقيق تحسن مماثل في أداء التصنيف. ومن هنا تبرز أهمية أساليب اختيار الخصائص في تحديد الخصائص الأكثر فائدة وتقليل أبعاد البيانات.

11. اختيار الخصائص وخوارزمياته المعتمدة في الدراسة

يُعد اختيار الخصائص إحدى الخطوات المهمة في معالجة البيانات ضمن تطبيقات التعلم الآلي، ويهدف إلى تحديد الخصائص الأكثر أهمية وارتباطًا بمتغير التصنيف، مع استبعاد الخصائص الأقل فائدة، بما يسهم في تقليل أبعاد البيانات وتبسيط النماذج وخفض المتطلبات الحسابية، مع المحافظة على أداء التصنيف قدر الإمكان.

وتُصنَّف أساليب اختيار الخصائص إلى ثلاثة أنواع رئيسية، هي: أساليب التصفية(Filter Methods) ، وأساليب التغليف (Wrapper Methods)، والأساليب المدمجة .(Embedded Methods) وتختلف هذه الأساليب من حيث آلية تقييم أهمية الخصائص؛ إذ تعتمد أساليب التصفية على مقاييس إحصائية أو معلوماتية مستقلة عن خوارزمية التصنيف، بينما تعتمد أساليب التغليف على أداء نموذج التعلم الآلي، في حين تنفذ الأساليب المدمجة عملية الاختيار ضمن مرحلة تدريب النموذج.

وفي هذه الدراسة، تم اعتماد خوارزميتين من أساليب التصفية، هما:

1. المعلومات المتبادلة :(Mutual Information – MI) تقيس مقدار المعلومات التي توفرها الخاصية حول الفئة المستهدفة، وتُستخدم لتقييم الخصائص وترتيبها وفقًا لمقدار المعلومات المتبادلة، ثم اختيار الخصائص الأعلى تقييمًا.
2. اختبار مربع كاي :(Chi-Square) يُستخدم لقياس مدى وجود علاقة إحصائية بين الخصائص والفئات المستهدفة، ويُعتمد على قيم الاختبار لترتيب الخصائص واختيار الأكثر ارتباطًا بعملية التصنيف.

وقد تم اختيار هاتين الخوارزميتين بهدف المقارنة بينهما من حيث الخصائص الناتجة ودرجة توافقها واستقرارها، إلى جانب تقييم أثر الخصائص المختارة في أداء نماذج التعلم الآلي وكفاءتها الحسابية عند تصنيف أنواع متعددة من الهجمات.

12. استقرار اختيار الخصائص ومقاييس تقييمه

لا يقتصر تقييم أساليب اختيار الخصائص على عدد الخصائص المحددة أو أداء نموذج التصنيف الناتج عنها، بل يشمل أيضًا مدى استقرار الخصائص المختارة عند تكرار عملية الاختيار باستخدام عينات بيانات مختلفة أو أساليب اختيار متعددة.

ويقصد باستقرار اختيار الخصائص مدى اتساق مجموعات الخصائص الناتجة عن تكرار عملية الاختيار في ظروف تجريبية محددة. فقد تختلف الخصائص المختارة عند تغيير عينة البيانات أو طريقة الاختيار، حتى عند استخدام العدد نفسه من الخصائص.

وتكتسب هذه المسألة أهمية في أنظمة كشف التسلل؛ إذ يساعد تحديد الخصائص التي تتكرر أهميتها عبر التجارب المختلفة في تقييم موثوقية أساليب الاختيار وقابلية تكرار نتائجها. ومع ذلك، فإن ارتفاع الاستقرار لا يعني بالضرورة ارتفاع دقة التصنيف، لذلك ينبغي تقييمه بالاقتران مع أداء النماذج وكفاءتها الحسابية.

ومن المقاييس المستخدمة لتقييم استقرار اختيار الخصائص معامل جاكارد(Jaccard Similarity) ، الذي يقيس نسبة عدد الخصائص المشتركة بين مجموعتين إلى عدد الخصائص الموجودة في اتحادهما، ويُحسب وفق المعادلة الآتية:

$$J(A,B)=\frac{A \cap B}{A \cup B}$$

حيث:

- (A):مجموعة الخصائص المختارة في التجربة الأولى.
- (B):مجموعة الخصائص المختارة في التجربة الثانية.
- ($|A \cap B|$):عدد الخصائص المشتركة بين المجموعتين.
- ($|A \cup B|$):عدد الخصائص الموجودة في المجموعتين دون تكرار.

تتراوح قيمة معامل جاكارد بين صفر وواحد؛ إذ تشير القيمة القريبة من الواحد إلى ارتفاع التشابه بين مجموعتي الخصائص، بينما تشير القيمة القريبة من الصفر إلى انخفاض التشابه. وعند إجراء عدة تجارب، يمكن حساب المعامل لكل زوج من مجموعات الخصائص، ثم حساب متوسط القيم الناتجة لتقدير مستوى الاستقرار العام.

وفي هذه الدراسة، يُستخدم معامل جاكارد في مستويين من التقييم: الأول لقياس مدى التوافق بين الخصائص المختارة بواسطة خوارزميتي MI وChi-Square، والثاني لتقييم استقرار الخصائص المختارة بواسطة خوارزمية MI عند تكرار الاختيار باستخدام عينات تدريب مختلفة. ويساعد ذلك على تحليل مدى اتساق نتائج الاختيار، وربطها بأداء نماذج كشف التسلل متعددة الفئات وكفاءتها الحسابية.

13.كفاءة اختيار الخصائص

تُعد كفاءة اختيار الخصائص من الجوانب المهمة في تقييم أساليب اختيار الخصائص المستخدمة في أنظمة كشف التسلل، إذ لا يقتصر التقييم على قدرتها على تحديد الخصائص الأكثر ملاءمة للتصنيف، بل يشمل أيضًا التكلفة الزمنية والحسابية لعملية الاختيار، ومدى تأثير الخصائص الناتجة في أداء نماذج التعلم الآلي.

وتشير كفاءة اختيار الخصائص إلى قدرة الأسلوب على تقليل عدد الخصائص المستخدمة في بناء نموذج التصنيف، مع المحافظة على مستوى مناسب من الأداء وتقليل المتطلبات الحسابية. وتزداد أهمية ذلك في أنظمة كشف التسلل نتيجة الحجم الكبير لبيانات حركة الشبكة وتعدد خصائصها، مما قد يؤدي إلى زيادة زمن المعالجة والتدريب عند استخدام جميع الخصائص المتاحة. ويمكن تقييم كفاءة أساليب اختيار الخصائص من خلال مجموعة من المؤشرات، أبرزها:

1. عدد الخصائص المختارة :(Number of Selected Features) يوضح مقدار تقليل أبعاد البيانات مقارنة بعدد الخصائص الأصلية، ويساعد على تحديد مدى الاختزال الذي يحققه أسلوب الاختيار.
2. زمن اختيار الخصائص :(Feature Selection Time) يمثل الوقت اللازم لتنفيذ خوارزمية اختيار الخصائص وتحديد المجموعة الناتجة، ويستخدم للمقارنة بين الأساليب من حيث التكلفة الزمنية لعملية الاختيار.
3. زمن التدريب :(Training Time) يشير إلى الوقت اللازم لتدريب نموذج التصنيف باستخدام الخصائص المختارة، ويساعد على تقييم أثر تقليل الخصائص في سرعة بناء النموذج.
4. زمن التنبؤ :(Prediction Time) يمثل الوقت اللازم للنموذج لتصنيف بيانات جديدة، وتبرز أهميته في التطبيقات التي تتطلب تحليل حركة الشبكة بسرعة مناسبة.
5. مؤشرات أداء التصنيف :تشمل الدقة(Accuracy) ، والإحكام(Precision) ، والاستدعاء(Recall) ، ومقياسF1-score ، ومعدل الإيجابيات الكاذبة(False Positive Rate – FPR) ، وتستخدم لتقييم قدرة النماذج على تصنيف الأنشطة الطبيعية والهجومية بصورة صحيحة.

وبناءً على ذلك، لا يمكن الحكم على كفاءة أسلوب اختيار الخصائص اعتمادًا على تقليل عدد الخصائص وحده، بل ينبغي مراعاة العلاقة بين مقدار الاختزال، والتكلفة الزمنية، وأداء التصنيف. فقد يؤدي تقليل الخصائص إلى خفض زمن التدريب والتنبؤ، لكنه قد يؤثر في قدرة النموذج على التمييز بين فئات الهجمات المختلفة.

وفي هذه الدراسة، تُقيَّم كفاءة أسلوبي MI و Chi-Squareمن خلال عدد الخصائص المختارة، وأثرها في زمن تدريب نماذج التعلم الآلي وزمن التنبؤ، إلى جانب مؤشرات أداء التصنيف متعدد الفئات، بهدف تحديد مدى تحقيق التوازن بين تقليل أبعاد البيانات وجودة التصنيف والكفاءة الحسابية.

14. كفاءة أنظمة كشف التسلل ومقاييس التقييم

1.14. مقاييس أداء التصنيف

تُعد مقاييس أداء التصنيف من المؤشرات الأساسية لتقييم كفاءة نماذج التعلم الآلي في أنظمة كشف التسلل، إذ تساعد على قياس قدرة النماذج على تصنيف حركة الشبكة بصورة صحيحة، واكتشاف فئات الهجمات المختلفة، وتقليل الإنذارات الخاطئة. وفي هذه الدراسة، تم اعتماد مجموعة من المقاييس الإحصائية لتقييم أداء التصنيف متعدد الفئات، وهي:

1. الدقة الكلية(Accuracy)

تقيس نسبة العينات التي صنفها النموذج تصنيفًا صحيحًا إلى إجمالي عدد العينات. وفي التصنيف متعدد الفئات، تُحسب وفق المعادلة:

$$\text{Accuracy} = \frac{-\sum_{i=1}^{K} \text{TPi}}{N}$$

حيث TP_i تمثل عدد العينات المصنفة بصورة صحيحة في الفئة i و N يمثل إجمالي عدد العينات.

2- متوسط الاحكام (Macro Precision)

يقيس متوسط الدقة مدى صحة التنبؤات التي أجراها النموذج لكل فئة، وذلك بحساب الدقة لكل فئة على حدة، ثم حساب المتوسط الحسابي لهذه القيم.

$$\text{Macro Precision} = \frac{1}{K}\sum_{i=1}^{K} \text{Precisioni}$$

حيث

$$\text{Precision} = \frac{TP}{TP+FP}$$

3- متوسط الاستدعاء(Macro Recall)

يقيس متوسط الاستدعاء قدرة النموذج على اكتشاف العينات التابعة لكل فئة بصورة صحيحة، ويُحسب من خلال المتوسط الحسابي لقيم الاستدعاء لجميع الفئات.

$$\text{Macr Recall} = = \frac{1}{K}\sum_{i=1}^{K} \text{Recalli}$$

حيث

$$\text{Recall} = \frac{TP}{TP+FN}$$

4- متوسط مقياس(Macro F1-Score) F1

يمثل هذا المقياس المتوسط الحسابي لقيم F1 الخاصة بجميع الفئات، ويجمع بين الدقة والاستدعاء، مما يوفر تقييمًا متوازنًا لأداء النموذج في التصنيف متعدد الفئات.

$$\text{Macro F1} = = \frac{1}{K}\sum_{i=1}^{K} \text{F1i}$$

حيث

$$\text{F1} = 2 * \frac{\text{Precisioni} * \text{Recalli}}{\text{Precisioni} + \text{Recalli}}$$

5- معدل الإيجابيات الكاذبة(False Positive Rate – FPR)

يقيس نسبة العينات التي تنتمي فعليًا إلى فئات سلبية، ولكن النموذج صنّفها خطأً على أنها تنتمي إلى فئة إيجابية. ويُعد هذا المقياس مهمًا في أنظمة كشف التسلل لتقييم مدى قدرة النموذج على تقليل الإنذارات الخاطئة.

$$\text{FPR}_i = \frac{FPi}{FPi+TNi}$$

ويُحسب متوسط معدل الإيجابيات الكاذبة لجميع الفئات وفق المعادلة التالية:

$$\text{Macro FPR} = \frac{1}{K}\sum_{i=1}^{K} \text{FPRi}$$

حيث يمثل انخفاض هذه القيمة قدرة أفضل على تقليل التصنيفات الخاطئة.

حيث

الإيجابيات الصحيحة(True Positive)	TP
السلبيات الصحيحة(True Negative)	TN
الإيجابيات الكاذبة(False Positive)	FP
السلبيات الكاذبة(False Negative)	FN
العدد الكلي للفئات	K
رقم الفئة	I
اجمالي عدد العينات	N

2.14. مقاييس الكفاءة الحسابية

إلى جانب مقاييس أداء التصنيف، تتطلب الدراسة تقييم التكلفة الحسابية الناتجة عن استخدام مجموعات الخصائص المختلفة، وذلك من خلال المؤشرات التالية:

1. عدد الخصائص(Number of Features)

يمثل عدد الخصائص المستخدمة في تدريب نموذج التصنيف، ويُستخدم لتحديد مقدار تقليل أبعاد البيانات بعد تطبيق أساليب اختيار الخصائص مقارنة باستخدام جميع الخصائص الأصلية.

2. زمن التدريب(Training Time)

يمثل الوقت المستغرق في تدريب نموذج التعلم الآلي باستخدام مجموعة الخصائص المحددة، ويُحسب من خلال الفرق بين وقت انتهاء التدريب ووقت بدايته:

$$T_{training} = T_{end} - T_{start}$$

3. زمن التنبؤ(Prediction Time)

يمثل الوقت اللازم للنموذج لإنتاج التصنيفات للعينات الجديدة، ويُستخدم لتقييم سرعة النموذج عند معالجة بيانات الاختبار.

وتساعد هذه المؤشرات مجتمعة على تحديد أثر تقليل الخصائص في المتطلبات الزمنية والحسابية لنماذج كشف التسلل.

3.14. العلاقة بين اختيار الخصائص وكفاءة أنظمة كشف التسلل

يسهم اختيار الخصائص في تقليل أبعاد البيانات وتخفيض المتطلبات الحسابية، إلا أن مقدار التحسن يعتمد على طريقة الاختيار وطبيعة الخصائص وخوارزمية التصنيف المستخدمة. فقد تحقق مجموعة مختصرة من الخصائص أداءً قريبًا من أداء المجموعة الأصلية مع انخفاض زمن التدريب والتنبؤ، بينما قد يؤدي استبعاد بعض الخصائص المهمة إلى تراجع القدرة على اكتشاف فئات معينة من الهجمات.

كما أن ارتفاع الدقة الكلية لا يعني بالضرورة تحقيق أداء متوازن بين جميع الفئات، خاصة عند التعامل مع مجموعات بيانات غير متوازنة. ولذلك، يتطلب تقييم كفاءة أنظمة كشف التسلل الجمع بين مقاييس التصنيف، ومقاييس الكفاءة الحسابية، ومقاييس استقرار اختيار الخصائص.

4.14. ارتباط مقاييس التقييم بأهداف الدراسة

ترتبط مقاييس التقييم المستخدمة في هذه الدراسة بأهدافها الرئيسة؛ إذ تُستخدم مقاييس أداء التصنيف لتقييم أثر طريقتي المعلومات المتبادلة (MI) ومربع كاي (Chi-Square) في قدرة نماذج التعلم الآلي على تصنيف حركة الشبكة إلى فئاتها المختلفة. كما تُستخدم مقاييس الكفاءة الحسابية لتحديد أثر تقليل عدد الخصائص في زمن التدريب والتنبؤ، بينما يُستخدم معامل جاكارد لتقييم التوافق بين مجموعات الخصائص المختارة، ومدى استقرار الاختيار عبر عينات التدريب المختلفة.

ويتيح الجمع بين هذه المقاييس إجراء مقارنة متكاملة بين أسلوبي اختيار الخصائص، وتحديد مدى قدرتهما على تحقيق التوازن بين استقرار الخصائص المختارة، وجودة التصنيف، وتقليل المتطلبات الحسابية لأنظمة كشف التسلل متعددة الفئات.

15.منهجية البحث والإجراءات العملية

1.15بيئة العمل والأدوات المستخدمة

نُفذت التجارب العملية في هذه الدراسة باستخدام لغة البرمجة Python ضمن بيئة Anaconda و Jupyter Notebook، لما توفره من أدوات ومكتبات متخصصة في تحليل البيانات وتطبيق خوارزميات التعلم الآلي واختيار الخصائص.

واستُخدمت مكتبة Pandas لمعالجة البيانات وتنظيمها، ومكتبة NumPy لتنفيذ العمليات العددية، ومكتبة Scikit-learn لتطبيق خوارزميات اختيار الخصائص وتدريب نماذج التصنيف وتقييم أدائها، بالإضافة إلى مكتبة XGBoost لتطبيق خوارزمية التعزيز المتدرج.

وأُجريت التجارب بهدف مقارنة أسلوبي اختيار الخصائص MI) و (Chi-Squareمن حيث مجموعات الخصائص الناتجة واستقرارها، وتأثير تقليل عدد الخصائص في أداء نماذج كشف التسلل متعددة الفئات وكفاءتها الحسابية. وشمل التقييم قياس أزمنة التدريب والتنبؤ ومؤشرات أداء التصنيف.

كما رُوعي الفصل بين بيانات التدريب والاختبار، بحيث تُستخدم بيانات التدريب في عملية اختيار الخصائص وتدريب النماذج، مع الاحتفاظ ببيانات الاختبار لتقييم الأداء النهائي، وذلك للحد من احتمالية تسرب البيانات.

2.15. وصف مجموعة البيانات

تعتمد هذه الدراسة على مجموعة البيانات UNSW-NB15، التي طُوّرت في مختبر الأمن السيبراني بجامعة نيو ساوث ويلز في كانبيرا، بهدف دعم الأبحاث المتعلقة بكشف التسلل وتحليل حركة مرور الشبكات. وتتضمن المجموعة أنشطة طبيعية وأنماطًا متعددة من الهجمات الإلكترونية، إضافة إلى مجموعة من الخصائص المستخرجة من حركة الشبكة وتسميات تحدد طبيعة النشاط المسجل.

تضم مجموعة البيانات تسع فئات من الهجمات، هي Fuzzers :و Analysisو Backdoorو DoSو Exploits و Genericو Reconnaissanceو Shellcodeو Worms، إلى جانب فئة الحركة الطبيعيةNormal ، وبذلك تتكون من عشر فئات للتصنيف متعدد الفئات.

اعتمدت الدراسة على ملفي التدريب والاختبار الأصليين، حيث يتكون ملف التدريب من 175,341 سجلًا، بينما يتكون ملف الاختبار من 82,332 سجلًا. وتم اعتماد الخاصية (attack_cat) متغيرًا مستهدفًا للتصنيف، نظرًا لاحتوائها على الفئات العشر التي تمثل أنواع الهجمات والحركة الطبيعية. ويوضح الجدول التالي توزيع سجلات فئات التصنيف في بيانات التدريب:

ر ت	الفئة	عدد السجلات
1	Normal	56,000
2	Generic	40,000
3	Exploits	33,393
4	Fuzzers	18,184
5	DoS	12,264
6	Reconnaissance	10,491
7	Analysis	2,000
8	Backdoor	1,746
9	Shellcode	1,133
10	Worms	130
	الإجمالي	**175,341**

يتضح من الجدول وجود تفاوت في توزيع العينات بين الفئات؛ إذ تمثل فئة Normal أكبر الفئات بعدد 56,000 سجل، وبنسبة 31.94% من إجمالي بيانات التدريب، بينما تمثل فئة Worms أصغر الفئات بعدد 130 سجلًا، أي ما يقارب 0.07%. كما تظهر فئات Analysis و Backdoorو Shellcodeتمثيلًا محدودًا مقارنة بالفئات الأكبر.

ويُعد هذا التفاوت من العوامل المهمة التي ينبغي مراعاتها عند تقييم النماذج، إذ قد تحقق بعض الخوارزميات دقة كلية مرتفعة دون أن تتمكن من تصنيف الفئات الأقل تمثيلًا بكفاءة. لذلك، اعتمدت الدراسة على مقاييس متعددة، تشمل Macro Precision و Macro Recallو Macro F1-score، إلى جانب مصفوفة الالتباس ومعدل الإيجابيات الكاذبة.

3.15 المعالجة المسبقة للبيانات

شملت مرحلة المعالجة المسبقة استبعاد الحقول غير المستخدمة في عملية التصنيف، وهي الحقول التعريفية وغير المستهدفة، مع اعتماد الخاصية (attack_cat) متغيرًا مستهدفًا. كما عولجت الخصائص الفئوية باستخدام الترميز الأحادي (One-Hot Encoding)، مع الإبقاء على الخصائص الرقمية.

وبعد استبعاد الحقول غير المستخدمة وترميز الخصائص الفئوية، نتجت مجموعة بيانات تتكون من 194 خاصية رقمية، استُخدمت مدخلًا لخوارزميات اختيار الخصائص ونماذج التعلم الآلي.

واحتُفظ بالتقسيم الأصلي لمجموعة البيانات إلى تدريب واختبار، حيث استُخدمت بيانات التدريب، التي تضم 175,341 سجلًا، في عمليات اختيار الخصائص وتدريب النماذج، بينما استُخدمت بيانات الاختبار، التي تضم 82,332 سجلًا، لتقييم أداء النماذج بعد التدريب.

جدول (2): ملخص تجهيز مجموعة البيانات

البيان	القيمة
سجلات التدريب	175,341
سجلات الاختبار	82,332
عدد الفئات المستهدفة	10
عدد الخصائص قبل الترميز	42
عدد الخصائص بعد الترميز	194
عدد الخصائص المختارة بواسطة MI	20
عدد الخصائص المختارة بواسطة Chi-Square	20

16.آلية اختيار الخصائص

اعتمدت الدراسة على أسلوبي التصفية(Filter Methods) ، من خلال تطبيق خوارزميتي المعلومات المتبادلة (Mutual Information – MI)واختبار مربع كاي(Chi-square) ، بهدف تقييم أهمية الخصائص واختيار الأكثر ارتباطًا بمتغير التصنيف.(attack_cat)

ولتقييم استقرار الاختيار، قُسمت بيانات التدريب إلى خمس عينات فرعية مختلفة، تمثل كل منها 20% من بيانات التدريب، بواقع 35,068 سجلًا لكل عينة، مع المحافظة على التوزيع النسبي لفئات التصنيف باستخدام أسلوب التقسيم الطبقي (Stratified Sampling).

طُبقت كل خوارزمية بصورة مستقلة على العينات الخمس، وحُسبت درجات الخصائص ورُتبت تنازليًا، ثم اختيرت أفضل 20 خاصية لكل عينة. وقيس استقرار الاختيار باستخدام معامل جاكارد(Jaccard Similarity) ، من خلال مقارنة مجموعات الخصائص المختارة بين أزواج العينات، وحساب متوسط التشابه.

كما أُجريت مقارنة بين مجموعتي الخصائص الناتجتين عن MI و Chi-squareلتحديد الخصائص المشتركة والمنفردة بينهما.

1.16. نتائج استقرار اختيار الخصائص

1.1.16. استقرار اختيار الخصائص

أظهرت نتائج تطبيق خوارزميتي MI و Chi-squareعلى العينات الخمس أن متوسط معامل جاكارد بلغ 1.0000 لكلتا الخوارزميتين، وفق النتائج المعروضة في الجدول (1). وتشير هذه القيمة إلى تطابق مجموعات الخصائص العشرين المختارة بين العينات الخمس لكل خوارزمية، مما يعكس استقرارًا كاملًا في الاختيار ضمن ظروف التجربة الحالية.

جدول (1): متوسط استقرار خوارزميات اختيار الخصائص

خوارزمية الاختيار	عدد العينات	عدد الخصائص	متوسط جاكارد
MI	5	20	1.0000
Chi-square	5	20	1.0000

وتشير هذه النتائج إلى أن تغيير العينة المستخدمة لم يؤثر في مجموعة الخصائص المختارة بواسطة أي من الخوارزميتين، ضمن حدود التجربة المنفذة.

2.1.16. مقارنة الخصائص المختارة بواسطة الخوارزميتين

بعد اختيار أفضل 20 خاصية باستخدام كل من MI وChi-square، أُجريت مقارنة بين المجموعتين لتحديد مدى الاتفاق بين الخوارزميتين.
أظهرت النتائج وجود 13 خاصية مشتركة بين المجموعتين، بينما انفردت كل خوارزمية باختيار 7 خصائص. وبلغ معامل جاكارد 0.4815، مما يشير إلى وجود تشابه جزئي بين نتائج الخوارزميتين، رغم اختلاف المعايير التي تعتمد عليها كل منهما في تقييم أهمية الخصائص.

جدول (2): مقارنة الخصائص المختارة بواسطة MI وChi-square

معيار المقارنة	النتيجة
عدد الخصائص المختارة بواسطة MI	20
عدد الخصائص المختارة بواسطة Chi-square	20
عدد الخصائص المشتركة	13
عدد الخصائص المنفردة لكل خوارزمية	7
معامل جاكارد	0.4815

الخصائص المشتركة بين الخوارزميتين:

1. ct_dst_src_ltm
2. dbytes
3. dinpkt
4. dload
5. dmean
6. dpkts
7. dttl
8. rate
9. sbytes
10. sinpkt
11. sload
12. smean
13. sttl

توضح النتائج أن الخوارزميتين تتفقان في تحديد مجموعة من الخصائص المهمة لحركة الشبكة، في حين تختلفان في اختيار بعض الخصائص الأخرى، وهو ما يعكس اختلاف الأساس الإحصائي والمعلوماتي لكل منهما.

17.قياس الكفاءة الزمنية لخوارزميات اختيار الخصائص

تم تقييم الكفاءة الزمنية لخوارزميتي المعلومات المتبادلة (Mutual Information – MI) واختبار مربع كاي-Chi) (Square، من خلال قياس الزمن اللازم لحساب درجات الخصائص واختيار أفضل 20 خاصية من إجمالي 194 خاصية، باستخدام عينة موحدة تضم 35,068 سجلًا. وأُجريت عشر عمليات قياس لكل خوارزمية، ثم حُسب متوسط زمن التنفيذ والانحراف المعياري.

أظهرت النتائج أن خوارزمية MI سجلت متوسط زمن تنفيذ بلغ 14.5754 ثانية، بانحراف معياري قدره 3.8853 ثانية، بينما سجلت خوارزمية Chi-Square متوسط زمن بلغ 0.1114 ثانية، بانحراف معياري قدره 0.2035 ثانية.

وتشير هذه النتائج إلى أن متوسط زمن تنفيذ Chi-Square كان أقل بنحو 130.8 مرة مقارنة بخوارزميةMI ، مما يوضح تفوقها من حيث سرعة تنفيذ عملية اختيار الخصائص ضمن ظروف التجربة الحالية. ومع ذلك، يُلاحظ وجود تفاوت في أزمنة التنفيذ، خصوصًا في خوارزميةChi-Square ، وهو ما يستدعي مراعاة الانحراف المعياري عند تفسير النتائج.

جدول (3): مقارنة الكفاءة الزمنية لخوارزميتي اختيار الخصائص

الخوارزمية	متوسط الزمن (ثانية)	الانحراف المعياري (ثانية)
MI	14.5754	3.8853
Chi-Square	0.1114	0.2035

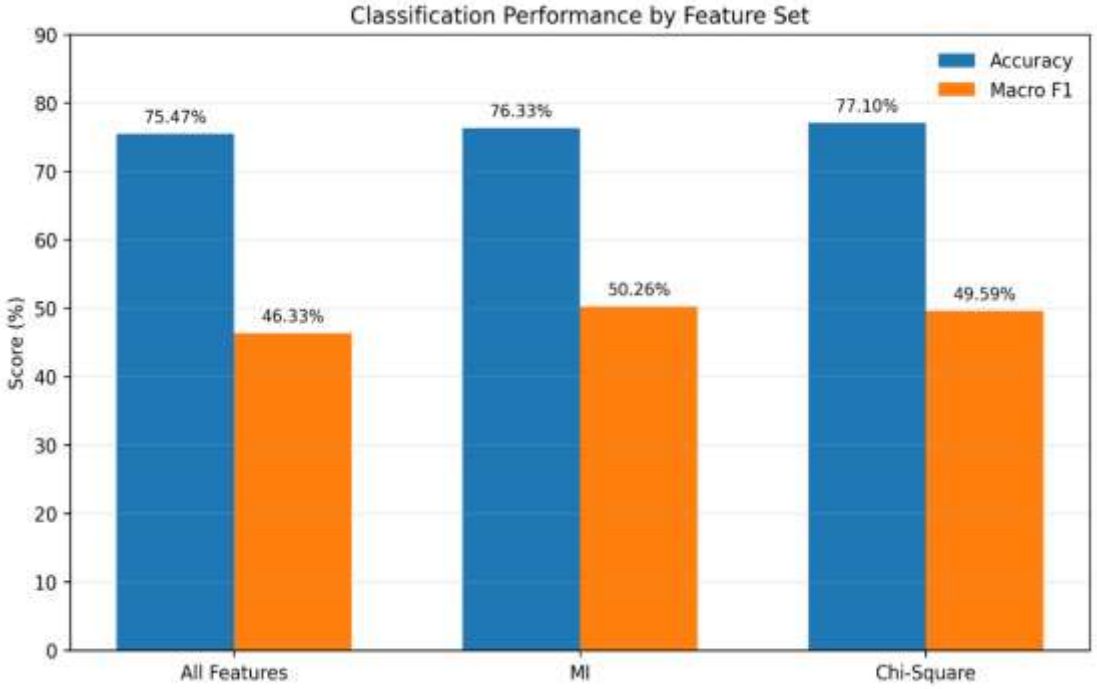
18.تقييم أثر الخصائص المختارة على أداء التصنيف متعدد الفئات

يهدف هذا الجزء إلى تقييم أثر اختيار الخصائص على أداء أنظمة كشف التسلل متعددة الفئات، من خلال مقارنة نتائج التصنيف باستخدام مجموعة الخصائص الكاملة (All Features) المكونة من 194 خاصية، ومجموعتي الخصائص المختارتين بواسطة خوارزميتي المعلومات المتبادلة (Mutual Information – MI) ومربع كاي(Chi-Square) ، بواقع 20 خاصية لكل مجموعة. وتم تقييم الأداء باستخدام مقاييس Accuracy و Macro Precisionو Macro Recallو Macro F1، بالإضافة إلى زمن التدريب، بهدف دراسة تأثير تقليل عدد الخصائص على جودة التصنيف والكفاءة الحسابية.

1.18. نتائج تقييم أداء التصنيف متعدد الفئات

أظهرت نتائج التجارب اختلافاً في أداء التصنيف باختلاف مجموعة الخصائص المستخدمة. فقد حقق النموذج باستخدام جميع الخصائص دقة كلية بلغت 75.47%، بينما ارتفعت إلى 76.33% عند استخدام خصائصMI ، وبلغت أعلى قيمة لها 77.10% عند استخدام خصائص.Chi-Square

في المقابل، حققت خوارزمية MI أفضل نتيجة وفق مقياس Macro F1 بقيمة 0.5026، مقارنة بـ0.4959 لخوارزمية Chi-Squareو0.4633 عند استخدام جميع الخصائص. وتشير هذه النتائج إلى أن اختيار الخصائص ساهم في تحسين أداء التصنيف، مع تفوق Chi-Square في الدقة الكلية، في حين حققت MI توازناً أفضل بين الدقة والاستدعاء عبر الفئات المختلفة.



2.18 تقييم الكفاءة الحسابية وزمن التدريب

أظهرت النتائج انخفاضاً ملحوظاً في زمن تدريب النموذج عند استخدام الخصائص المختارة مقارنة بمجموعة الخصائص الكاملة. فقد بلغ زمن التدريب 682.19 ثانية باستخدام جميع الخصائص، وانخفض إلى 471.20 ثانية باستخدامMI ، بينما سجلت Chi-Square أقل زمن تدريب بلغ 303.93 ثانية.

وبذلك حققت MI انخفاضاً في زمن التدريب بنسبة 30.9%، في حين بلغ الانخفاض 55.5% باستخدامChi-Square ، مقارنة بمجموعة الخصائص الكاملة. وتعكس هذه النتائج أن تقليل عدد الخصائص من 194 إلى 20 خاصية أسهم في تحسين الكفاءة الحسابية وتقليل الوقت اللازم للتدريب، مع المحافظة على أداء تصنيفي أفضل من استخدام جميع الخصائص.



3.18.تقييم استقرار اختيار الخصائص باستخدام معامل جاكارد

تم تقييم استقرار خوارزميتي المعلومات المتبادلة (Mutual Information – MI) واختبار كاي تربيع (Chi-Square) باستخدام خمس عينات عشوائية، تمثل كل منها 20% من بيانات التدريب، وباستخدام البذور العشوائية (10، 20، 30، 40، 50). وتم اختيار أفضل 20 خاصية لكل عينة، ثم حساب معامل جاكارد بين جميع أزواج العينات، بإجمالي عشر مقارنات لكل خوارزمية.

أظهرت النتائج أن خوارزمية MI حققت متوسط معامل جاكارد بلغ 1.0000، مما يشير إلى ثبات كامل في الخصائص المختارة عبر العينات الخمس. في المقابل، حققت خوارزمية Chi-Square متوسطاً بلغ 0.8883، مما يعكس مستوى استقرار مرتفعاً، مع وجود اختلافات محدودة في الخصائص المختارة بين العينات.

خوارزمية اختيار الخصائص	متوسط Jaccard	مستوى الاستقرار
MI	1.0000	كامل
Chi-Square	0.8883	مرتفع

تشير هذه النتائج إلى أن خوارزمية MI كانت أكثر استقراراً من Chi-Square في اختيار الخصائص، وفقاً للتجارب المنفذة على مجموعة بيانات.UNSW-NB15

4.18.تقييم أثر اختيار الخصائص على أداء التصنيف متعدد الفئات

أظهرت نتائج التجارب أن استخدام خوارزميتي المعلومات المتبادلة (Mutual Information – MI) واختبار كاي تربيع (Chi-Square) لاختيار أفضل 20 خاصية من أصل 194 خاصية أسهم في المحافظة على مستوى دقة التصنيف، مع تحقيق تحسن في بعض مقاييس الأداء مقارنة باستخدام جميع الخصائص.

حققت خوارزمية MI أعلى قيمة للدقة (Accuracy) بلغت 74.87%، كما سجلت أعلى قيمتي الاستدعاء (Recall) ومقياسF1-score ، حيث بلغتا 56.57% و52.26% على التوالي. في المقابل، حققت مجموعة جميع الخصائص أعلى قيمة للدقة الإيجابية (Precision) بنسبة 53.22.%

ومن ناحية الكفاءة الحسابية، حققت خوارزمية Chi-Square أفضل أداء من حيث زمن التنفيذ، إذ بلغ زمن التدريب 326.42 ثانية، وزمن التنبؤ 1.09 ثانية، مقارنة بـ586.16 ثانية و4.69 ثانية عند استخدام جميع الخصائص.

تشير هذه النتائج إلى أن تقليل عدد الخصائص يمكن أن يحقق توازناً بين أداء التصنيف والكفاءة الحسابية، مع اختلاف مزايا كل خوارزمية؛ إذ أظهرت MI أداءً أفضل وفق مقياسي Recall وF1-score، بينما تميزت Chi-Square بسرعة التدريب والتنبؤ.

جدول: نتائج تقييم أداء التصنيف متعدد الفئات باستخدام مجموعات الخصائص المختلفة

مجموعة الخصائص	عدد الخصائص	Accuracy	Precision	Recall	F1-score
All Features	194	0.7477	0.5322	0.5147	0.4844
MI	20	0.7487	0.5293	0.5657	0.5226
Chi-Square	20	0.7486	0.5063	0.5528	0.5108

يمكن استخلاص أن خوارزمية MI كانت الأفضل من حيث جودة التصنيف وفق مقياسF1-score ، في حين كانت Chi-Squareالأفضل من حيث الكفاءة الزمنية. وبالتالي، فإن اختيار الخوارزمية الأنسب يعتمد على متطلبات نظام كشف التسلل، سواء كانت الأولوية لتحسين اكتشاف الهجمات أو تقليل زمن المعالجة.

تشير النتائج إلى أن استخدام خوارزميات اختيار الخصائص يمكن أن يسهم في تحسين أداء التصنيف متعدد الفئات وتقليل التكلفة الحسابية للتدريب. وقد أظهرت خوارزمية MI أفضل قيمة لمقياسMacro F1 ، مما يجعلها أكثر ملاءمة عند الاهتمام بتوازن أداء التصنيف بين الفئات، بينما حققت خوارزمية Chi-Square أفضل دقة كلية وأقل زمن تدريب.

وبناءً على ذلك، لا يمكن تحديد أفضل خوارزمية اعتماداً على مقياس واحد فقط، بل ينبغي مراعاة طبيعة المقياس المستهدف ومتطلبات نظام كشف التسلل. وتؤكد النتائج أهمية تقييم خوارزميات اختيار الخصائص من حيث جودة التصنيف والكفاءة الحسابية معاً.

5.18. تقييم الكفاءة الحسابية لخوارزميات اختيار الخصائص

تم تقييم الكفاءة الحسابية لخوارزميتي المعلومات المتبادلة (Mutual Information – MI) واختبار كاي تربيع (Chi-Square)من خلال قياس زمن التدريب وزمن التنبؤ لنماذج التصنيف باستخدام مجموعات الخصائص المختلفة، ومقارنة النتائج مع استخدام جميع الخصائص.(All Features)

أظهرت النتائج أن استخدام 20 خاصية بدلاً من 194 خاصية أسهم في تقليل الزمن اللازم لتدريب نماذج التصنيف والتنبؤ. فقد حققت خوارزمية Chi-Square أقل زمن تدريب بلغ 326.42 ثانية، مقارنة بـ477.42 ثانية لخوارزمية MI و586.16 ثانية عند استخدام جميع الخصائص.

كما سجلت Chi-Square أقل زمن للتنبؤ بلغ 1.09 ثانية، مقابل 3.40 ثانية لخوارزمية MI و4.69 ثانية لجميع الخصائص.

جدول: مقارنة الكفاءة الحسابية باستخدام مجموعات الخصائص المختلفة

مجموعة الخصائص	عدد الخصائص	زمن التدريب (ثانية)	زمن التنبؤ (ثانية)
All Features	194	586.1609	4.6886
MI	20	477.4248	3.4042
Chi-Square	20	326.4193	1.0886

وتشير النتائج إلى أن استخدام خوارزمية MI أدى إلى تقليل زمن التدريب بنسبة 18.55% وزمن التنبؤ بنسبة 27.39% مقارنة باستخدام جميع الخصائص. في حين حققت خوارزمية Chi-Square انخفاضاً أكبر بلغ 44.31% في زمن التدريب و76.78% في زمن التنبؤ.

وبناءً على ذلك، أظهرت خوارزمية Chi-Square كفاءة حسابية أعلى من خوارزميةMI ، من حيث تقليل زمن التدريب والتنبؤ، مما يشير إلى أهمية اختيار الخصائص في تحسين كفاءة تنفيذ أنظمة كشف التسلل متعددة الفئات.

19.الاستنتاجات

استناداً إلى التجارب والنتائج التي تم الحصول عليها باستخدام مجموعة بياناتUNSW-NB15 ، توصلت الدراسة إلى الاستنتاجات التالية:

1. أظهرت نتائج اختيار الخصائص أن استخدام خوارزميتي المعلومات المتبادلة (MI) واختبار كاي تربيع-Chi) (Squareأسهم في تقليل عدد الخصائص من 194 إلى 20 خاصية، مع المحافظة على مستوى متقارب من دقة التصنيف متعدد الفئات.
2. حققت خوارزمية MI أعلى مستوى من استقرار اختيار الخصائص، بمتوسط معامل جاكارد بلغ 1.0000 عبر العينات الخمس، مقارنة بـ0.8883 لخوارزميةChi-Square ، مما يشير إلى ثبات أكبر في الخصائص المختارة باستخدام MIضمن التجارب المنفذة.
3. أظهرت المقارنة بين الخصائص المختارة بواسطة الخوارزميتين وجود 13 خاصية مشتركة من أصل 27 خاصية في اتحاد المجموعتين، بمعامل جاكارد بلغ 0.4815، مما يعكس اختلافاً في الخصائص التي تحددها كل خوارزمية وفقاً لمعيار الاختيار الخاص بها.
4. حققت خوارزمية MI أفضل أداء في التصنيف وفق مقياسي Recall وMacro F1-score، حيث بلغت قيمتاهما 0.5657 و0.5226 على التوالي، كما سجلت أعلى دقة كلية (Accuracy) بلغت 0.7487.
5. أظهرت خوارزمية Chi-Square كفاءة حسابية أعلى من حيث زمن تدريب نموذج التصنيف وزمن التنبؤ، حيث سجلت أقل زمن تدريب بلغ 326.42 ثانية، وأقل زمن تنبؤ بلغ 1.09 ثانية.
6. تشير النتائج إلى أن تقييم خوارزميات اختيار الخصائص لا ينبغي أن يعتمد على دقة التصنيف وحدها، بل يجب أن يشمل استقرار الخصائص المختارة والكفاءة الحسابية، نظراً لاختلاف أداء الخوارزميتين باختلاف معيار التقييم.

وبناءً على ذلك، تُظهر الدراسة أن خوارزمية MI توفر توازناً أفضل من حيث استقرار اختيار الخصائص وجودة التصنيف وفق مقياسF1-score ، بينما تتميز خوارزمية Chi-Square بكفاءة زمنية أعلى، مما يجعل اختيار الخوارزمية المناسبة مرتبطاً بمتطلبات نظام كشف التسلل وأولوياته.

20. الخاتمة

تناولت هذه الدراسة تقييم استقرار وكفاءة خوارزميات اختيار الخصائص في أنظمة كشف التسلل متعددة الفئات، من خلال مقارنة خوارزميتي المعلومات المتبادلة (MI) واختبار كاي تربيع (Chi-Square) باستخدام مجموعة بيانات UNSW-NB15.

وأوضحت النتائج أن اختيار الخصائص يمثل خطوة مهمة في تقليل أبعاد البيانات وتحسين كفاءة نماذج التصنيف، مع المحافظة على مستوى مناسب من الأداء. كما بينت الدراسة أن استقرار الخصائص المختارة يختلف باختلاف خوارزمية الاختيار، وأن ارتفاع الاستقرار لا يعني بالضرورة تحقيق أفضل أداء في جميع مقاييس التصنيف.

وقد أظهرت خوارزمية MI تميزاً في استقرار الاختيار وأداء التصنيف وفق مقياس Macro F1 ، بينما حققت خوارزمية Chi-Square كفاءة زمنية أفضل من حيث التدريب والتنبؤ. ويؤكد ذلك أن المفاضلة بين خوارزميات اختيار الخصائص ينبغي أن تستند إلى مجموعة متكاملة من المعايير، وليس إلى دقة التصنيف وحدها.

وتوصي الدراسة بتوسيع نطاق التقييم مستقبلاً ليشمل مجموعات بيانات أخرى، واختبار خوارزميات إضافية لاختيار الخصائص، ودراسة مدى استقرار النتائج عند استخدام نماذج تصنيف وبيئات بيانات مختلفة، بما يسهم في تطوير أنظمة كشف تسلل أكثر كفاءة وموثوقية.

21. المراجع

1. More, S., Idrissi, M., Mahmoud, H., & Asyhari, A. T. (2024). Enhanced intrusion detection systems performance with UNSW-NB15 data analysis. Algorithms, 17(2), 64. https://doi.org/10.3390/a17020064
2. Sowmya, T., & Mary Anita, E. A. (2025). A novel stable feature selection algorithm for machine learning based intrusion detection system. Procedia Computer Science, 252, 738–747. https://doi.org/10.1016/j.procs.2025.01.034
3. Umer, M., Tahir, M., Sardaraz, M., Sharif, M., Elmannai, H., & Algarni, A. D. (2025). Network intrusion detection model using wrapper based feature selection and multi head attention transformers. Scientific Reports, 15, Article 28718. https://doi.org/10.1038/s41598-025-11348-5
4. Saputra, A., Ramli, K., Nugroho, A. S., Nugraha, I. G. D., & Pranggono, B. (2026). A novel hybrid IGL1 feature selection method for high-performance intrusion detection on the UNSW-NB15 dataset using multiple machine learning models. Big Data and Cognitive Computing, 10(6), 182. https://doi.org/10.3390/bdcc10060182
5. Khraisat, A., Gondal, I., Vamplew, P., & Kamruzzaman, J. (2019). Survey of intrusion detection systems: techniques, datasets and challenges. Cybersecurity, 2, Article 20. https://doi.org/10.1186/s42400-019-0038-7.
6. Dina, A. S., & Manivannan, D. (2021). Intrusion detection based on Machine Learning techniques in computer networks. Internet of Things, 16, 100462. https://doi.org/10.1016/j.iot.2021.100462.
7. Moustafa, N., & Slay, J. (2015). UNSW-NB15: A comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set). In *2015 Military Communications and Information Systems Conference (MilCIS)* (pp. 1–6). IEEE. https://doi.org/10.1109/MilCIS.2015.7348942