الجامعة السورية الخاصة
كلية هندسة الذكاء الصنعي
قسم هندسة أمن نظم المعلومات

مشروع أعد لنيل شهادة الهندسة في تخصص أمن نظم المعلومات والشبكات الحاسوبية (فصلي)

كشف السلوك الشاذ في حركة البيانات الخاصة بشبكة انترنت الأشياء باستخدام الذكاء الاصطناعي

Anomaly detection for IOT network traffic using graphical interfacing in real time

إعداد الطلاب

سعود عبدالعزيز الحجي عبيدة محمد لامع فتوح

إشراف

د. وسيم جنيدي

تشرين الاول 2025

جدول المحتويات

جدول الاشكال

جدول الجداول

الخلاصة

المقدمة

في عصر التحول الرقمي المتسارع، أصبحت أجهزة إنترنت الأشياء (IoT) تشكل العمود الفقري للبنية التحتية الذكية في مختلف القطاعات، بدءًا من المنازل الذكية ووصولًا إلى المدن الذكية والصناعات المتطورة. ومع التوسع الهائل في استخدام هذه الأجهزة، ازدادت أهميتها كعناصر أساسية في تحسين الكفاءة التشغيلية وتقديم خدمات ذكية. إلا أن هذا الانتشار الواسع رافقَه تزايد ملحوظ في حجم التهديدات الأمنية التي تستهدف هذه الأجهزة، بسبب ضعف آليات الحماية المضمنة فيها، وعدم توفر أنظمة مراقبة ذكية قادرة على كشف الأنشطة الضارة في الوقت المناسب.

يهتم تخصص "أمن نظم المعلومات الذكية" بضمان استمرارية عمل هذه الأنظمة، وحماية البيانات المنقولة عبر الشبكات، وتأمين الموارد الحساسة من الوصول غير المصرح به، والتصدي للهجمات التي قد تؤدي إلى تعطيل الخدمات أو خسارة المعلومات. في هذا الإطار، تُعد أنظمة كشف الشذوذ وأدوات تحليل حركة الشبكة في الزمن الحقيقي من الأدوات الحيوية التي تسهم في تعزيز منظومة الدفاع الأمني للمؤسسات.

من بين أنواع الهجمات السيبرانية التي تستهدف شبكات إنترنت الأشياء، تبرز هجمات الحرمان من الخدمة(DDoS) ، وهجمات الانتحال، والوصول غير المصرح به. تتم هذه الهجمات عادةً باستغلال نقاط ضعف في البروتوكولات أو نظام التحكم، مما يسمح للمهاجم بالتحكم في الأجهزة أو تعطيلها. تشير الأدلة إلى أن ضعف آليات المصادقة، وعدم تحديث البرامج الثابتة، وعدم تفعيل مبدأ "الأقل امتيازًا"، تزيد من قابلية التعرض لمثل هذه الهجمات. وبالتالي، فإن التصدي لهذه التهديدات يمثل ضرورة أمنية ملحة لضمان استقرار واستمرارية الخدمات المقدمة عبر هذه الأجهزة.

على الرغم من توفر أدوات مراقبة الشبكات وتحليل الأحداث، إلا أن العديد من المؤسسات لا تزال تعاني من صعوبة في ربط الأنشطة الشبكية المشبوهة بالأحداث المخزنة في قواعد البيانات وأنظمة السجلات. بعبارة أخرى: رغم وجود أنظمة SIEM قادرة على جمع وتحليل السجلات، فإن الكشف الاستباقي عن الهجمات في بيئات إنترنت الأشياء ليس دائمًا شاملاً. كما أن تحليل النشاط في الزمن الحقيقي واتخاذ إجراءات تلقائية ما يزال يشكل تحديًا في البيئات الحقيقية. من هنا تنشأ الحاجة إلى تصميم نظام كشف شذوذ مخصص لمراقبة حركة بيانات أجهزة إنترنت الأشياء، يعمل على الربط بين تنبيهات الشبكة وتحليل الأمن لضمان استجابة أسرع ودقة أعلى في الكشف.

الهدف الرئيسي

يهدف هذا البحث إلى تصميم وتنفيذ نظام كشف شذوذ مخصص لاكتشاف الهجمات في شبكات إنترنت الأشياء، ضمن بيئة محلية تحاكي الشبكات الذكية، ويعمل في الزمن الحقيقي. كما يهدف إلى تعزيز قدرة المؤسسات على اكتشاف محاولات الاختراق والأنشطة غير الطبيعية، وربط التنبيهات المستخلصة من حركة الشبكة بقواعد البيانات المؤسسية لتحليل مركزي وتحقيق استجابة أوتوماتيكية.

يعالج هذا المشروع الفجوة البحثية التي تشير إلى أن أنظمة كشف الشذوذ التقليدية غالبًا ما تكون غير ملائمة لبيئات إنترنت الأشياء بسبب طبيعة الأجهزة محدودة الموارد، وتنوع البروتوكولات المستخدمة. كما يستند البحث إلى تطورات حديثة في تقنيات الكشف المعتمدة على تحليل حركة البيانات الصادرة باستخدام النمذجة السلوكية، حيث أظهرت دراسات سابقة نسبة دقة تجاوزت 97% في كشف هجمات DDoS في شبكات IoT.

كما يتطلع البحث إلى مواكبة التطور السريع في أساليب الهجوم المعقدة والمتغيرة، ويقترح استخدام منهجية قائمة على القواعد مع إمكانية التوسع مستقبلاً لدمج تقنيات التعلم الآلي لزيادة فعالية الكشف وتقليل التنبيهات الكاذبة.

وبالتالي، فإن النتائج المتوقعة من هذا المشروع تشمل: تحسين دقة كشف الهجمات في شبكاتIoT ، وتقليل زمن الاستجابة للحوادث الأمنية، وتوفير نموذج عملي قابل للتطبيق في سياقات مؤسسية متنوعة.

التطبيقات العملية

يمكن تطبيق هذا النظام عمليًا في عدة بيئات ومجالات تشغيلية تعتمد على بنى تحتية ذكية، مثل:

- المؤسسات المالية التي تستخدم أجهزة استشعار ذكية لمراقبة الفروع.
- المنازل والمدن الذكية حيث يتم ربط الأجهزة عبر شبكات.IoT
- المراكز الأمنية التي تحتاج إلى مراقبة مستمرة للأنشطة غير الطبيعية.

على سبيل المثال:
في مؤسسة تستخدم كاميرات ذكية وأجهزة استشعار، يمكن للنظام مراقبة حركة البيانات الواردة والصادرة، وإرسال تنبيهات إلى لوحة تحكم عند رصد نمط يشير إلى هجوم محتمل مثل DDoS أو محاولة اختراق.

يمكن أيضًا دمج النظام مباشرة مع منصة مراقبة أمنية) مثل (SIEM لعرض مؤشرات مثل: عدد محاولات الهجوم خلال آخر 24 ساعة، وعناوين IP مصدر الهجوم، أو نوع الهجوم المكتشف.

إضافة إلى ذلك، يمكن تنفيذ استجابة تلقائية، مثل: عزل الجهاز المخترق، أو حظر عنوان IP مهاجم، أو إشعار مشرف الأمن عبر البريد الإلكتروني. كما أن النظام قابل للتوسع ليعمل في بيئات موزعة، حيث يتم جمع بيانات المراقبة من عدة أجهزة في منصة مركزية واحدة لتحليل موحد، مما يعزز قدرة المؤسسة على التعامل مع هجمات منسقة أو موزعة.

التحديات

يواجه تنفيذ هذا النظام عدة تحديات تقنية ومنهجية، منها:

1. تنوع وتطور هجمات إنترنت الأشياء:
لا تقتصر الهجمات على نمط محدد، بل تشمل أنواعًا متعددة مثل هجماتDDoS ، وهجمات العبث بالبيانات، والهجمات الخفية، مما يجعل الاعتماد على قواعد ثابتة فقط غير كافٍ.

2. الدمج مع أنظمة المراقبة المركزية:(SIEM)
يجب ألا يقتصر عمل النظام على الكشف المحلي فحسب، بل ينبغي أن ينقل التنبيهات إلى منصة SIEM لتحليلها مع أحداث أخرى (مثل محاولات تسجيل دخول فاشلة، أو تغييرات في إعدادات الأجهزة). هذا يتطلب تنسيقًا بين الأدوات، وضبط بروتوكولات الإرسال، وتوحيد تنسيق السجلات.

3. تقليل التنبيهات الكاذبة وتحسين صيانة النظام:

تحتاج قواعد الكشف إلى تحديث مستمر لمواكبة الهجمات الجديدة. كما أن ارتفاع معدل التنبيهات الكاذبة قد يؤدي إلى إهمال التنبيهات الحقيقية. أظهرت دراسات سابقة أن عدم دقة قواعد الكشف يضعف من كفاءة النظام بشكل ملحوظ.

الخاتمة

في الختام، يمثل تصميم نظام متخصص لكشف الشذوذ في شبكات إنترنت الأشياء خطوة حيوية نحو تعزيز الأمن السيبراني في ظل التوسع المتزايد للبنى التحتية الذكية. يواجه هذا المجال تحديات فريدة، نابعة من طبيعة الأجهزة محدودة الموارد وتنوع البروتوكولات واتساع نطاق الهجمات المتطورة، مما يجعل الحلول التقليدية غير كافية.

يهدف هذا البحث إلى سد فجوة مهمة من خلال اقتراح نظام يعمل في الزمن الحقيقي، يعتمد على نهج قائم على القواعد قابل للتكيف والدمج مع الأنظمة المركزية مثل .SIEM من المتوقع أن يساهم هذا النظام في رفع دقة الاكتشاف، وتقليل زمن الاستجابة للحوادث، وخفض نسبة التنبيهات الكاذبة، مما يقدم حلاً عملياً يمكن تطبيقه في قطاعات متنوعة كالمدن الذكية والمؤسسات المالية والمراكز الأمنية.

على الرغم من التحديات المتعلقة بتنوع الهجمات وتعقيد عمليات الدمج والصيانة، فإن التطور المستمر للتقنيات، مع إمكانية دمج التعلم الآلي في المستقبل، يبشر ببناء أنظمة دفاع أكثر مرونة وذكاءً. يعزز هذا النهج الشمولي قدرة المؤسسات على حساسية بنيتها التحتية الرقمية، وضمان استمرارية الخدمات، والحفاظ على سلامة البيانات في عالم يتجه نحو الارتباط الدائم لكل شيء.

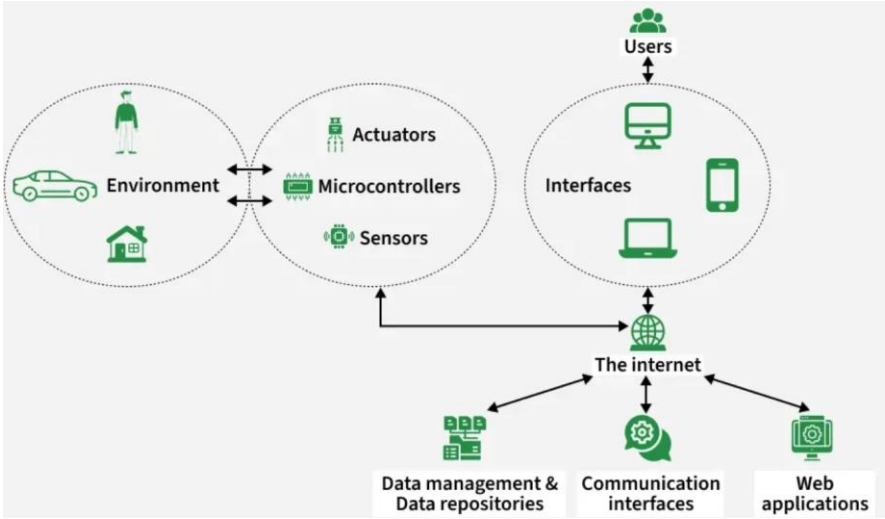
الفصل الأول

الدراسة النظرية

1.1 الخلفية العلمية والنظرية

1.1.1 إنترنت الأشياء: الثورة الرقمية الجديدة والتحديات المصاحبة

في عصر التحول الرقمي المتسارع، تشكل أجهزة إنترنت الأشياء (IoT) العمود الفقري للبنية التحتية الذكية المعاصرة، حيث امتدت تطبيقاتها من المنازل الذكية إلى المدن الذكية، والرعاية الصحية، والصناعة الذكية. هذا الانتشار الواسع رافقه نمو هائل في عدد الأجهزة المتصلة، حيث تشير التقديرات إلى وصول عددها إلى عشرات المليارات بحلول عام 2025. ومع هذا النمو، تبرز تحديات جوهرية في مجال الأمن السيبراني، نابعة من الطبيعة الفريدة لهذه البيئات التي تتسم بتنوع البروتوكولات، واختلاف إمكانيات الأجهزة، ومحدودية الموارد الحاسوبية والطاقية. هذه الخصائص تجعل من الصعب تطبيق حلول الأمن التقليدية التي تعتمد على قواعد ثابتة وتوقيعات محددة مسبقاً.



رسم توضيحي 1 كيف تتصل أجهزة إنترنت الأشياء عبر الشبكات السلكية

1.1.2 أمن إنترنت الأشياء: من النهج التقليدي إلى الذكاء الاصطناعي

تواجه أنظمة أمن إنترنت الأشياء التقليدية فجوة متسعة بين قدراتها ومتطلبات البيئات الحديثة، حيث أصبحت الهجمات السيبرانية أكثر تطوراً وذكاءً، قادرة على التمويه والتحور بشكل ديناميكي. الهجمات الحديثة مثل الهجمات السلوكية (Behavioral Attacks) وهجمات الانزياح التدريجي (Slow-Drift Attacks) تتسم بقدرتها على محاكاة السلوك الطبيعي للأجهزة، مما يجعل اكتشافها باستخدام القواعد الثابتة مهمة شبه مستحيلة. في هذا السياق، يبرز الذكاء الاصطناعي والتعلم الآلي كحلول ثورية قادرة على تحليل الأنماط المعقدة، والتعلم من البيانات التاريخية، والكشف عن الانحرافات الدقيقة التي قد تشير إلى هجمات مخفية. هذا التحول من الكشف التفاعلي إلى الكشف التنبؤي والاستباقي يمثل نقلة نوعية في فلسفة تأمين بيئات إنترنت الأشياء.

1.2 المشكلة العلمية: الفجوة بين التعقيد والقدرات

على الرغم من التطور الكبير في تقنيات الأمن السيبراني، تواجه أنظمة إنترنت الأشياء إشكالية جوهرية تتمثل في صعوبة كشف السلوك الشاذ في حركة البيانات (Network Traffic) في الوقت الحقيقي. هذه الصعوبة تنبع من تعقيدين متزامنين: أولاً، الطبيعة الديناميكية والمتطورة

للتهديدات السيبرانية التي أصبحت قادرة على تجاوز آليات الكشف التقليدية. ثانياً، القيود الصارمة التي تفرضها أجهزة IoT من حيث استهلاك الطاقة، وقدرة المعالجة، والعرض الترددي المتاح.

تشير الدراسات الحديثة إلى أن أكثر من 70% من أجهزة IoT تفتقر إلى آليات أمنية كافية للكشف عن الهجمات في الوقت الفعلي، مما يؤدي إلى زيادة فترة الكشف عن الاختراقات التي قد تصل إلى عدة أشهر في بعض الحالات. هذا التأخير في الكشف يزيد من حجم الخسائر المادية والمعنوية، ويجعل الأنظمة عرضة لتهديدات متطورة مثل هجمات DDoS الموزعة والهجمات الخفية.(Stealthy Attacks)

الإشكالية الأساسية التي يواجهها الباحثون والمطورون تكمن في التوفيق بين متطلبات الكفاءة في استخدام الموارد المحدودة من ناحية، وضرورة تحقيق دقة عالية في الكشف مع تقليل الإنذارات الكاذبة من ناحية أخرى. الحلول التقليدية لكشف الشذوذ غالباً ما تكون ثقيلة جداً على أجهزة IOT محدودة الموارد، بينما الحلول الخفيفة تفتقر إلى الدقة الكافية للتعامل مع التعقيد المتزايد للتهديدات.

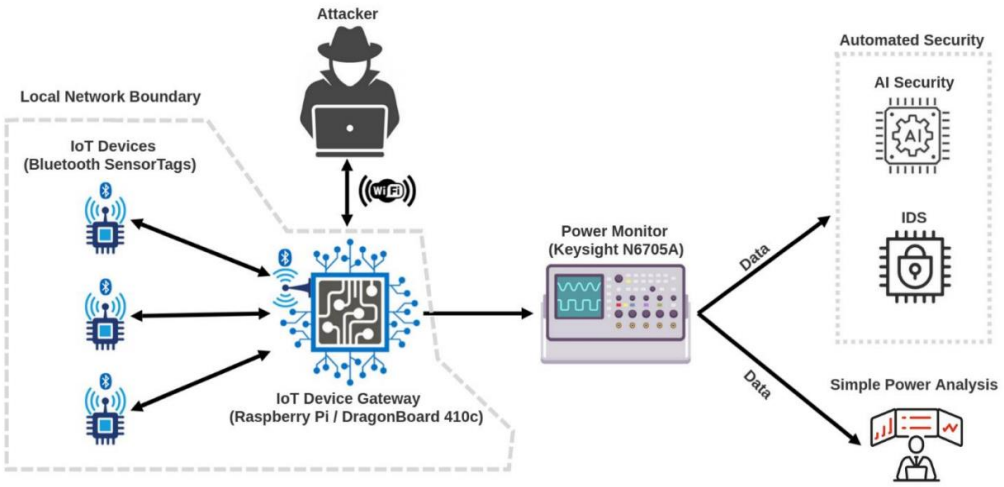
1.3 الهدف من البحث: نحو نموذج ذكي متكامل

يسعى هذا البحث إلى تطوير إطار عمل متكامل لكشف الشذوذ في حركة بيانات IoT يعتمد على أحدث تقنيات الذكاء الاصطناعي والتعلم الآلي، مع مراعاة القيود الفريدة لبيئات إنترنت الأشياء. يهدف البحث إلى تحقيق مجموعة من الأهداف المتكاملة:

أولاً، تصميم نموذج كشف شذوذ ذكي وخفيف الوزن (Intelligent Lightweight Anomaly Detection Model) يمكن تشغيله على أجهزة ذات إمكانيات محدودة، مع الحفاظ على كفاءة الأداء من حيث استهلاك الطاقة والذاكرة. يعتمد هذا النموذج على تقنيات التعلم العميق الخفيفة (Lightweight Deep Learning) القادرة على تحليل الأنماط المعقدة في حركة البيانات.

ثانياً، تطوير منهجية هجينة تجمع بين مزايا التعلم الآلي المتقدم وتحليل السلوكيات(Behavioral Analysis) ، مع إمكانية التكيف الديناميكي مع الأنماط الجديدة للهجمات. يتضمن ذلك استخدام تقنيات مثل الشبكات العصبية التلافيفية (CNNs) للتعرف على الأنماط المكانية، والشبكات العصبية المتكررة (RNNs/LSTMs) لتحليل التسلسلات الزمنية.

ثالثاً، إنشاء بيئة اختبار واقعية (Testbed) تحاكي شبكة IOT حقيقية، تتضمن مجموعة متنوعة من الأجهزة الذكية، لتقييم أداء النموذج في ظروف تشغيلية قريبة من الواقع. هذه البيئة تسمح باختبار قدرة النموذج على التعامل مع تنوع البروتوكولات واختلاف أنماط حركة البيانات.



رسم توضيحي 2 بيئة اختبار أو مراقبة محلية

رابعاً، إجراء مقارنة منهجية مع النماذج المرجعية في الأدبيات الحديثة، لتحديد مزايا وعيوب النموذج المقترح، وتقديم توصيات عملية قابلة للتطبيق لتحسين أداء أنظمة كشف الشذوذ في بيئات.IoT

1.4 الدراسات المرجعية: التطور نحو الذكاء الاصطناعي

تم النظر في 10 دراسات مرجعية عامة حول المشروع عامة , بالإضافة الى 5 دراسات مرجعية خاصة حول الخوارزميات التي تم استخدامها ضمن المشروع

الدراسة (العنوان)	السنة	المحتوى / المساهمة	العلاقة بمشروعنا
Recent advances in anomaly detection in Internet of Things	2024	تحليل شامل لأحدث التطورات في تقنيات كشف الشذوذ الخاصة ببيئاتIoT ، مع التركيز على تقنيات التعلم الآلي الخفيفة والتحليل السلوكي للشبكة.	توفر خريطة طريق واضحة للتقنيات الناشئة وتساعد في تحديد الأساليب الواعدة التي يمكن تبنيها أو تطويرها في مشروعنا.
A systematic review of anomaly detection in IoT security	2024	مراجعة منهجية تحدد الفجوات البحثية في المنهجيات الحالية، مثل نقص المعايير الموحدة للتقييم وندرة الأطر الديناميكية.	تساعد في استهداف مشكلات بحثية حقيقية وغير مستهلكة، وتوجه تصميم نظامنا نحو معالجة هذه الفجوات.
Enhancing anomaly detection and prevention in Internet of Things	2025	نهج هجين يجمع بين الشبكات العصبية العميقة (DNNs)وتقنية البلوكشين لتأمين سجلات الأنشطة والتنبيهات.	يفتح آفاقاً للتوسع المستقبلي، خاصة في تعزيز موثوقية وسلامة نظام الإبلاغ عن الحوادث في مشروعنا.
A secure framework for the Internet of Things anomalies	2025	إطار عمل متكامل يدمج مراحل جمع البيانات، المعالجة المسبقة، الكشف باستخدام خوارزميات تعلم آلي، والاستجابة.	يوفر نموذجاً عملياً يمكن محاكاته في تصميم نظامنا، كما يشكل معياراً أدائياً للمقارنة.
An anomaly detection system for data quality assurance in IoT	2024	استخدام تقنيات التعلم غير المراقب (Unsupervised Learning)لضمان جودة البيانات في أجهزة الاستشعار.	نموذج بديل ممتاز للمقارنة مع نهجنا القائم على تحليل حركة المرور، ويساعد في تقييم فعالية كل منهجية.

A Literature Review on Security in the Internet of Things	2024	مراجعة شاملة لطبقات الأمان المختلفة (الفيزيائية، الشبكية، التطبيقية) في بيئات.IoT	يساعد في وضع نظام كشف الشذوذ في سياقه الصحيح ضمن هيكل دفاعي متعدد الطبقات.
A Comparative Analysis of Anomaly Detection Methods in IoT Networks	2025	تحليل مقارن كمي لأداء خوارزميات متعددة (Isolation Forest, LSTM, Autoencoders)باستخدام معايير أداء موحدة.	دليل عملي لاختيار الأنسب من الخوارزميات لبيئة مشروعنا بناءً على مقايضات الأداء والكفاءة.
Current research on Internet of Things security protocols	2025	تحليل نقاط القوة والضعف في بروتوكولات الاتصال الشائعة (MQTT, CoAP, DTLS) في بيئات.IoT	يساعد في توقع ناقلات الهجوم المحتملة وتطوير قواعد كشف أكثر دقة تستهدف ثغرات بروتوكولات محددة.
Securing the Connected World: A Review Paper of IoT Security	2024	تنظيم التحديات والحلول الحديثة في إطار واحد يوفر رؤية شاملة للمجال.	يضمن أن مشروعنا لا يعمل بمعزل عن الجهود البحثية الأوسع، بل يكون متوافقاً مع الاتجاهات العامة.
Generative AI for Internet of Things Security: Challenges and Opportunities	2025	استكشاف تطبيقات الذكاء الاصطناعي التوليدي في أمن IoT، مثل توليد بيانات هجومية واقعية لتدريب الأنظمة.	يحدد مساراً واضحاً للتطوير المستقبلي والابتكار في مشروعنا، خاصة في مراحل التطوير المتقدمة.

جدول1 الدراسات المرجعية في أمن إنترنت الأشياء

الخوارزمية	الدراسة (العنوان)	السنة	المحتوى / المساهمة	العلاقة بمشروعنا
SVM (Support Vector Machine)	A Hybrid Intrusion Detection System Based on SVM for IoT Networks	2023	تقديم نظام هجين لكشف التسلل يعتمد على SVMمع تحسين النواة (Kernel) للتعامل مع بيانات شبكات IoT عالية الأبعاد.	توفر أساساً قوياً لتنفيذ SVM محسّن في مشروعنا، مع التركيز على ضبط معاملات النواة لتحسين الأداء في البيئات محدودة الموارد.
Isolation Forest	Lightweight Anomaly Detection for IoT Using Improved Isolation Forest	2024	تطوير نسخة محسنة من Isolation Forest بخوارزمية عزل متقدمة تقلل من تعقيد الحساب مع الحفاظ على دقة الكشف في بيانات IOT المتوازنة.	مباشرة قابلة للتطبيق في مشروعنا، حيث توفر حلاً عملياً للتعامل مع البيانات غير المتوازنة مع الحفاظ على كفاءة الموارد.
Density-Based Spatial Clustering (DBSCAN)	Real-time Anomaly Detection in IoT Using Adaptive DBSCAN	2023	تقديم نسخة متكيفة من DBSCAN تقوم بضبط معاملات التجمع minPts و ε تلقائياً بناءً على كثافة بيانات الشبكة في الوقت الحقيقي.	توفر منهجية متقدمة للكشف عن الهجمات الجماعية(Burst Attacks) في مشروعنا، مع إمكانية التكيف مع التغيرات في أنماط حركة البيانات.

جدول2 الدراسات المرجعية في خوارزميات الذكاء الاصطناعي المستخدمة

1.4.1 تحليل الدراسات المرجعية وتكاملها مع المشروع

تشكل الدراسات المرجعية المذكورة في الجدولين أعلاه الأساس المعرفي المتين الذي يبنى عليه هذا المشروع البحثي. من خلال التحليل المتعمق لهذه الدراسات، يمكن استخلاص عدة اتجاهات رئيسية:

- أولاً :هناك تحول واضح في مجال أمن إنترنت الأشياء من الحلول التقليدية القائمة على القواعد الثابتة نحو أنظمة ذكية تعتمد على التعلم الآلي والتعلم العميق. الدراسات الحديثة (2024-2025) تركّز بشكل متزايد على تطوير نماذج خفيفة الوزن (Lightweight Models)يمكن نشرها على أجهزة IoT محدودة الموارد.

- ثانياً :تبرز أهمية النهج الهجينة (Hybrid Approaches) التي تجمع بين مزايا خوارزميات متعددة. دراسة "Ensemble Learning for IoT Security" (2024) تظهر أن الجمع بين SVM و Isolation Forest يمكن أن يحقق دقة أعلى ومعدل إنذارات كاذبة أقل مقارنة باستخدام كل خوارزمية بشكل منفصل.
- ثالثاً :تشير الدراسات إلى أن التعامل مع البيانات غير المتوازنة (Imbalanced Data) يمثل تحديًا رئيسيًا في تدريب نماذج الذكاء الاصطناعي لأمن IoT. دراسة "Lightweight Anomaly Detection for IoT Using Improved Isolation Forest" (2024) تقدم حلاً عملياً لهذا التحدي من خلال تحسين خوارزمية Isolation Forest للتعامل مع البيانات التي يكون فيها عدد العينات الهجومية أقل بكثير من العينات الطبيعية.
- رابعاً :تكشف المراجعات المنهجية مثل "A systematic review of anomaly detection in IoT security" (2025) عن فجوات بحثية مهمة، أبرزها نقص المعايير الموحدة لتقييم أداء أنظمة الكشف، وندرة النماذج القابلة للتكيف الديناميكي مع تطور أنماط الهجوم.
- خامساً :الدراسات المتعلقة بالخوارزميات المحددة(SVM ، Isolation Forest، DBSCAN) تقدم رؤى عملية حول كيفية تحسين هذه الخوارزميات لتتناسب مع بيئات IoT. على سبيل المثال، دراسة SVM للكشف عن التسلل (2023) توضح أهمية اختيار نواة مناسبة (Kernel Selection) لتحسين قدرة الخوارزمية على فصل البيانات عالية الأبعاد.

هذه الرؤى المستخلصة من الدراسات المرجعية توجه تصميم وتنفيذ مشروعنا الحالي، حيث نسعى إلى تطوير نظام كشف شذوذ يجمع بين مزايا الخوارزميات المذكورة، مع معالجة الفجوات البحثية التي حددتها الدراسات الحديثة، وتصميم نموذج خفيف الوزن يتناسب مع القيود الفريدة لبيئات إنترنت الأشياء.

1.5 الخلاصة والفجوات البحثية

1.5.1 استنتاجات الفصل الأول

يؤسس هذا الفصل الركيزة النظرية والأكاديمية المتينة التي يبنى عليها التصميم العملي والتجريبي في هذا المشروع البحثي. من خلال التحليل الشامل الذي قدمه هذا الفصل، يمكن استخلاص عدة استنتاجات أساسية:

أولاً، أكد التحليل أن بيئات إنترنت الأشياء (IoT) تمتلك طبيعة فريدة تفرض تحديات غير مسبوقة على آليات الأمن التقليدية. هذه الطبيعة تتجلى في ثلاثة أبعاد رئيسية :**القيود الحاسوبية والطاقية** الصارمة التي تمنع تشغيل أنظمة الأمن التقليدية الثقيلة، **تنوع البروتوكولات** وعدم توحيد المعايير الذي يعقّد عملية المراقبة والتحليل، و**الحساسية للكفاءة** في استخدام النطاق الترددي الذي يجعل حلول الأمن التقليدية غير عملية

ثانياً، أظهر التحليل أن التهديدات الأمنية التي تستهدف شبكات IoT قد تطورت بشكل ملحوظ، حيث انتقلت من الهجمات البسيطة إلى هجمات أكثر ذكاءً وتعقيداً .**هجمات الشذوذ (Anomaly Attacks)** تحديداً تبرز كتهديد رئيسي نظراً لقدرتها على التمويه والاندماج مع السلوك الطبيعي للشبكة. هذه الهجمات، مثل الهجمات السلوكية (Behavioral Attacks) وهجمات الانزياح التدريجي (Slow-Drift Attacks)، تستغل الثغرات في آليات المصادقة والمراقبة في الأجهزة محدودة الموارد

ثالثاً، أوضحت الدراسة أن الاعتماد على حلول الأمن التقليدية القائمة على التوقيعات (Signature-based) لم يعد كافياً في مواجهة التحديات المعاصرة .**التعقيد الديناميكي** للتهديدات الحديثة يتطلب الانتقال إلى نماذج كشف أكثر ذكاءً وقدرة على التكيف والتعلم من البيانات. هذا ما يفسر التحول الواضح في الأبحاث الحديثة نحو اعتماد تقنيات الذكاء الاصطناعي والتعلم الآلي

رابعاً، بينت المقارنة بين أنواع أنظمة كشف الشذوذ (Anomaly Detection Systems) أهمية **تبني نهج هجين (Hybrid Approach)** يوازن بين الكفاءة في استخدام الموارد والدقة في الكشف. الأنظمة القائمة على الحافة (Edge-based) توفر استجابة سريعة وحماية للخصوصية ولكنها محدودة الموارد، بينما الأنظمة المعتمدة على السحابة (Cloud-based) تقدم قدرة معالجة عالية ولكنها تعتمد على اتصال مستمر وتواجه تأخيرات في الاستجابة الجمع بين مزايا النهجين يبدو الأمثل لبيئات IoT.

خامساً، كشفت الدراسة المرجعية عن **فجوات بحثية مهمة** تحتاج إلى معالجة. أبرز هذه الفجوات تشمل: نقص المعايير الموحدة لتقييم أداء أنظمة الكشف، محدودية النماذج القابلة للتكيف الديناميكي مع تطور أنماط الهجوم، وندرة الأطر الشاملة التي تجمع بين الكفاءة في استخدام الموارد والدقة العالية في الكشف

1.5.2 الفجوات البحثية التي يعالجها المشروع

بناءً على التحليل السابق، يمكن تحديد الفجوات البحثية الرئيسية التي يسعى هذا المشروع إلى معالجتها:

- فجوة الكفاءة والدقة في النماذج الخفيفة :بينما توجد نماذج خفيفة الوزن مناسبة لأجهزة IoT ، فإنها غالباً ما تفتقر إلى الدقة الكافية للتعامل مع التعقيد المتزايد للتهديدات. من ناحية أخرى، النماذج عالية الدقة تكون ثقيلة جداً على الموارد المحدودة. يسعى هذا المشروع إلى سد هذه الفجوة من خلال تطوير نموذج يجمع بين خفة الوزن ودقة الكشف.
- فجوة التكيف الديناميكي :معظم أنظمة كشف الشذوذ الحالية تعتمد على نماذج ثابتة يتم تدريبها مسبقاً، مما يحد من قدرتها على التكيف مع أنماط هجمات جديدة غير مرئية أثناء التدريب يهدف هذا المشروع إلى تطوير نموذج ذو قدرة تكيفية، يمكن تحديثه وتطويره بشكل شبه ذاتي مع ظهور تهديدات جديدة.
- فجوة التقييم الموحد :تفتقر الأبحاث الحالية إلى معايير موحدة وشاملة لتقييم أداء أنظمة كشف الشذوذ في بيئات IoT غالباً ما تركز الدراسات على مقاييس دقة تقليدية) مثل (Accuracy وتهمل مقاييس الأداء في الزمن الحقيقي وكفاءة استخدام الموارد. يسعى هذا المشروع إلى تطوير إطار تقييم متعدد الأبعاد يشمل جميع الجوانب الحرجة.
- فجوة التكامل العملي :هناك فصل بين الأبحاث الأكاديمية والتطبيقات العملية في مجال أمن .IoT العديد من النماذج المقدمة في الأبحاث تكون معقدة جداً أو تتطلب موارد غير متوفرة في البيئات الحقيقية. يهدف هذا المشروع إلى تقديم نموذج عملي قابل للتطبيق في بيئات تشغيلية حقيقية.
- فجوة معالجة البيانات غير المتوازنة: البيانات الأمنية في بيئات IoT تتسم بعدم التوازن الشديد، حيث تكون العينات الطبيعية أكثر بكثير من العينات الهجومية معظم الخوارزميات التقليدية تفشل في التعامل مع هذا التحدي. يسعى المشروع إلى تطوير تقنيات متقدمة لمعالجة عدم التوازن في البيانات.

1.5.3 مساهمة المشروع المتوقعة

يتوقع أن يساهم هذا المشروع في سد الفجوات المذكورة من خلال:

- نظرياً :تقديم إطار نظري متكامل يجمع بين أحدث تقنيات الذكاء الاصطناعي ومتطلبات بيئات IoT المحددة، مع تطوير منهجية جديدة لتصميم نماذج خفيفة الوزن وعالية الدقة.
- منهجياً :تطوير إطار تقييم شمولي متعدد الأبعاد يشمل مقاييس الدقة التقليدية، ومقاييس الأداء في الزمن الحقيقي، ومقاييس كفاءة استخدام الموارد، مما يوفر معياراً موحداً للمقارنة بين الأنظمة المختلفة.
- عملياً :تصميم نموذج كشف شذوذ قابل للتطبيق في بيئات IoT حقيقية، مع وثائق تفصيلية للتنفيذ والتكامل مع الأنظمة القائمة.
- تطبيقياً :تقديم توصيات عملية للمؤسسات والجهات المعنية حول كيفية تحسين أمن شبكات IoT باستخدام التقنيات المقترحة.

الفصل الثاني

المنهجية والتجهيز

2.1 المقدمة المنهجية : نحو بيئة ذكية للتجريب

يهدف هذا الفصل إلى وصف المنهجية الشاملة والمتكاملة التي اعتمدها البحث لتطوير وتقييم نموذج كشف الشذوذ المعتمد على الذكاء الاصطناعي لبيئات إنترنت الأشياء. لا يقتصر الأمر على مجرد تجميع للأجهزة والبرمجيات، بل يشمل تأسيس منهجية علمية قابلة للتكرار تسمح بإجراء تجارب عادلة وقابلة للمقارنة. تم تصميم البيئة التجريبية لمحاكاة التحديات الحقيقية التي تم تحليلها في الفصل السابق، بما في ذلك محدودية النطاق الترددي، وقيود المعالجة، وتنوع البروتوكولات، مع إضافة بُعد جديد هو قدرة النظام على التعلم والتكيف باستمرار.

2.2 مراحل العمل : مسار متكامل من البيانات إلى الذكاء

اعتمد البحث على مسار منهجي متعدد المراحل يضمن التطور المتدرج والمتسق للنموذج المقترح:

A. المرحلة الأولى: إعداد قاعدة البيانات : تم الحصول على مجموعة بيانات شاملة وحديثة تمثل بيئة IoT حقيقية، مع تطبيق تقنيات معالجة بيانات متقدمة لتحضيرها للتدريب على نماذج الذكاء الاصطناعي.

B. المرحلة الثانية: التصميم والتدريب الذكي : تطوير وتدريب نماذج تعلم آلي وتعليم عميق متعددة، مع تحسين معاملاتها بناءً على أدائها في كشف الأنماط المعقدة للهجمات.

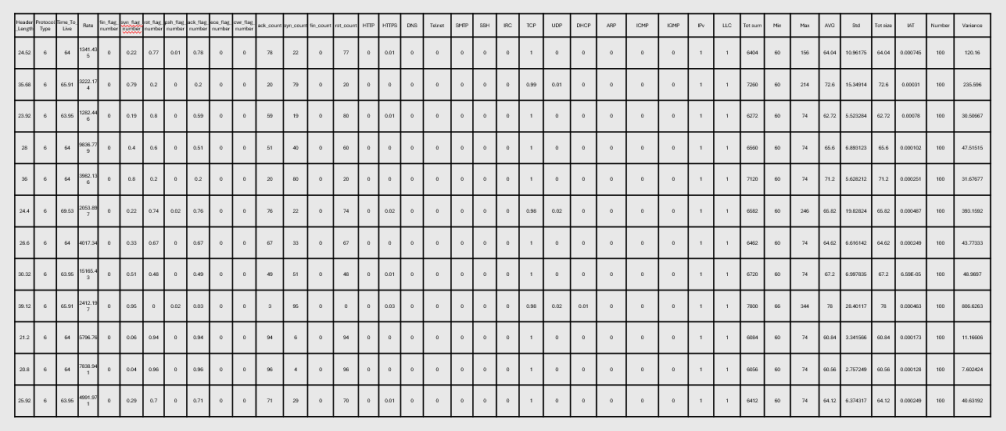
C. المرحلة الثالثة: المحاكاة الواقعية : إنشاء بيئة شبكية تحاكي شبكة IoT حقيقية، تسمح باختبار النماذج في ظروف تشغيلية مشابهة للواقع.

D. المرحلة الرابعة: التكامل والاختبار الديناميكي : دمج النماذج المدربة في البيئة المحاكاة، واختبار قدرتها على الكشف في الوقت الحقيقي، مع إجراء تحسينات تكرارية بناءً على الأداء الملاحظ.

2.3 مجموعة البيانات: أساس التدريب الذكي

2.3.1 مصدر البيانات واختياره

تم اختيار مجموعة البيانات CIC-IoT-2023 التي طورها مركز الاتصالات والحوسبة الكندي (Canadian Institute for Cybersecurity) كأساس للتدريب والتقييم. يعود اختيار هذه المجموعة إلى عدة اعتبارات منهجية: أولاً، كونها من أحدث مجموعات البيانات في مجال أمن إنترنت الأشياء، حيث تم تطويرها عام 2023 مما يضمن مواكبتها لأحدث أنواع الهجمات والتقنيات. ثانياً، شموليتها حيث تحتوي على بيانات شبكة تم جمعها من بيئة IoT حقيقية تشمل 105 جهازاً من 33 نوعاً مختلفاً، مما يوفر تنوعاً كبيراً في أنماط حركة البيانات. ثالثاً، دقة التصنيف حيث تضم مجموعة واسعة من الهجمات المصنفة بدقة، تتراوح من هجمات DDoS الموزعة إلى الهجمات الخفية والهجمات السلوكية.



Header_Length	Protocol_Type	Time_To_Live	Rate	fin_flag_number	syn_flag_number	rst_flag_number	psh_flag_number	ack_flag_number	ece_flag_number	cwr_flag_number	ack_count	syn_count	fin_count	rst_count	HTTP	HTTPS	DNS	Telnet	SMTP	SSH	IRC	TCP	UDP	DHCP	ARP	ICMP	IGMP	IPv	LLC	Tot sum	Min	Max	AVG	Std	Tot size	IAT	Number	Variance
24.52	6	64	1041.43 5	0	0.22	0.77	0.01	0.78	0	0	78	22	0	77	0	0.01	0	0	0	0	0	1	0	0	0	0	0	1	1	6404	60	156	64.04	10.96175	64.04	0.000745	100	120.16
35.68	6	65.91	1022.17 4	0	0.79	0.2	0	0.2	0	0	20	79	0	20	0	0	0	0	0	0	0	0.99	0.01	0	0	0	0	1	1	7280	60	214	72.6	15.34914	72.6	0.00031	100	235.596
23.52	6	63.55	1282.48 6	0	0.19	0.8	0	0.59	0	0	59	19	0	80	0	0.01	0	0	0	0	0	1	0	0	0	0	0	1	1	6272	60	74	62.72	5.525264	62.72	0.00076	100	30.50967
28	6	64	9636.77 9	0	0.4	0.6	0	0.51	0	0	51	40	0	60	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	6660	60	74	66.6	6.893123	66.6	0.000102	100	47.51515
96	6	64	3962.13 6	0	0.8	0.2	0	0.2	0	0	20	80	0	20	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	7120	60	74	71.2	5.628212	71.2	0.000251	100	31.67677
24.4	6	68.53	2653.88 7	0	0.22	0.74	0.02	0.76	0	0	76	22	0	74	0	0.02	0	0	0	0	0	0.98	0.02	0	0	0	0	1	1	9382	60	246	93.82	19.82824	93.82	0.000487	100	393.1592
28.6	6	64	4617.34	0	0.33	0.67	0	0.67	0	0	67	33	0	67	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	6462	60	74	64.62	6.690142	64.62	0.000249	100	43.77333
30.32	6	63.95	10165.4 9	0	0.51	0.48	0	0.49	0	0	49	51	0	48	0	0.01	0	0	0	0	0	1	0	0	0	0	0	1	1	6720	60	74	67.2	6.997836	67.2	6.59E-05	100	48.9697
38.12	6	65.91	2412.19 7	0	0.95	0	0.02	0.03	0	0	3	95	0	0	0	0.03	0	0	0	0	0	0.98	0.02	0.01	0	0	0	1	1	7800	60	344	78	28.40117	78	0.000403	100	806.8283
21.2	6	64	5796.76	0	0.06	0.94	0	0.94	0	0	94	6	0	94	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	6684	60	74	66.84	3.341566	66.84	0.000173	100	11.16606
20.8	6	64	7838.04 1	0	0.04	0.96	0	0.96	0	0	96	4	0	96	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	6656	60	74	66.56	2.757249	66.56	0.000124	100	7.602424
26.92	6	63.96	4604.07 1	0	0.29	0.7	0	0.71	0	0	71	29	0	70	0	0.01	0	0	0	0	0	1	0	0	0	0	0	1	1	6412	60	74	64.12	6.374317	64.12	0.000249	100	40.63192

رسم توضيحي 3 قاعدة البيانات للمشروع

2.3.2 معالجة البيانات للذكاء الاصطناعي

لمعالجة البيانات وتحضيرها للتدريب على نماذج الذكاء الاصطناعي، تم تطبيق سلسلة من العمليات المتقدمة:

التنظيف والتحقق : تم تطبيق خوارزميات كشف القيم الشاذة وإزالة الضوضاء لضمان جودة البيانات المدخلة للنماذج.

التطبيع والتحجيم : تم تطبيع القيم الرقمية باستخدام تقنيات مثل Min-Max Scaling و Z-score Normalization لضمان تقارب توزيعات البيانات.

معالجة عدم التوازن : نظراً للطبيعة غير المتوازنة لمجموعات بيانات الأمن السيبراني (حيث تكون العينات الطبيعية أكثر بكثير من العينات الهجومية)، تم استخدام تقنيات متقدمة مثل SMOTE (Synthetic Minority Over-sampling Technique) و ADASYN لإنشاء عينات هجومية اصطناعية متوازنة.

استخراج الميزات الذكية : تم تطوير خوارزميات لاستخراج ميزات متقدمة من البيانات الأولية، تشمل ميزات زمنية (مثل متوسط الفاصل الزمني بين الحزم)، وميزات إحصائية (مثل الانحراف المعياري لمعدل نقل البيانات)، وميزات متعلقة بالبروتوكولات مأنماط أعلامTCP

2.4 النماذج المقترحة: من التقليدية إلى الذكاء المتقدم

2.4.1 معايير اختيار النماذج

تم اختيار النماذج بناءً على مجموعة معايير شاملة تلبي متطلبات بيئات IoT المحدودة المواقع:

الكفاءة الحسابية : قدرة النموذج على العمل بكفاءة على أجهزة ذات إمكانيات محدودة.

القدرة على التعلم المعقد : قدرة النموذج على التعرف على الأنماط غير الخطية والمعقدة في بيانات الشبكة.

التكيف الديناميكي : إمكانية تحديث النموذج وتكيفه مع أنماط هجمات جديدة.

سرعة الاستنتاج : قدرة النموذج على إجراء التصنيف في زمن مناسب للتطبيقات الزمن الحقيقي.

القدرة التفسيرية : درجة فهم سلوك النموذج وقراراته، وهي مهمة للتطبيقات الأمنية الحساسة.

2.4.2 عائلة النماذج المقترحة

النماذج التقليدية المحسنة:

- Random Forest المحسن : مع تحسين معاملات عمق الشجرة وعدد المقدرات.
- XGBoost المتقدم : باستخدام تقنيات التعلم التدرجي مع معالجة فعالة للقيم المفقودة.
- **SVM المتطور مع النواة الذكية : تأخذ في الاعتبار الخصائص الفريدة لبيانات شبكات IOT**
- **DBSCAN المتكيف للتجميع الزمني-مكاني : للتعامل مع الطبيعة الزمنية-المكانية لبيانات شبكات IoT**

نماذج التعلم العميق الخفيفة:

- LSTM (Long Short-Term Memory) خفيف الوزن : مصمم خصيصاً لتحليل التسلسلات الزمنية في حركة الشبكة.
- CNN أحادي البعد : للتعرف على الأنماط المكانية في ميزات حركة البيانات.
- النموذج الهجين CNN-LSTM : يجمع بين مزايا كلا النموذجين لتحليل الأنماط الزمنية والمكانية معاً.

نماذج مبتكرة للكشف عن الشذوذ:

- Isolation Forest المتطور : مع تحسين معاملات العزل للتعامل مع البيانات عالية الأبعاد.
- Autoencoder خفيف الوزن : للكشف عن الانحرافات في أنماط حركة البيانات الطبيعية.

2.5 منهجية التقييم الشاملة

2.5.1 مقاييس الأداء المتعددة الأبعاد

تم تطوير إطار تقييم متعدد الأبعاد يشمل:

مقاييس الدقة التقليدية:

- الدقة(Accuracy) ، الاستدعاء(Recall) ، الدقة(Precision) ، ودرجة F1.
- منحنى ROC ومنطقة تحته (AUC-ROC).

مقاييس الأداء في الزمن الحقيقي:

- وقت الاستجابة (Response Time) الوقت من اكتشاف الحدث إلى إصدار التنبيه.
- وقت المعالجة (Processing Time) الوقت اللازم لتحليل دفعة من البيانات.

مقاييس كفاءة الموارد:

- استهلاك الذاكرة (Memory Footprint) مقدار الذاكرة المطلوبة لتشغيل النموذج.
- استهلاك الطاقة (Power Consumption) الطاقة المستهلكة أثناء التشغيل (محسوبة نظرياً أو عملياً).

مقاييس الجدوى العملية:

- معدل الإنذارات الكاذبة (False Positive Rate) وتأثيره على جدوى النظام العملي.

- قابلية التوسع (Scalability) قدرة النظام على التعامل مع أعداد متزايدة من الأجهزة.

2.5.2 منهجية التقييم المقارن

تم تصميم منهجية مقارنة منهجية تشمل:

- تقييم كل نموذج على مجموعة الاختبار المستقلة.
- مقارنة أداء النماذج المختلفة باستخدام نفس معايير التقييم.
- تحليل المقايضات (Trade-offs) بين الدقة وكفاءة الموارد.
- تقييم القدرة على التعميم (Generalization) على بيانات غير مرئية.

2.6 التحديات المنهجية والحلول المقترحة

2.6.1 تحديات معالجة البيانات الضخمة

واجه البحث تحدياً في معالجة الملايين من سجلات البيانات التي تحتوي عليها مجموعة .CIC-IoT-2023 تم التغلب على هذا التحدي من خلال:

- استخدام تقنيات المعالجة الدفعية.(Batch Processing)
- تطبيق خوارزميات أخذ العينات العشوائية الذكية.
- استخدام مكتبات معالجة بيانات متوازية مثل Dask و.Modin

2.6.2 تحدي التوازن بين الدقة والكفاءة

تم معالجة هذا التحدي من خلال:

- تطوير نماذج خفيفة الوزن باستخدام تقنيات مثل Pruning و.Quantization
- استخدام معماريات شبكية متخصصة للبيانات المتسلسلة.
- تطبيق تقنيات Transfer Learning لنماذج مسبقة التدريب.

2.6.3 تحدي التكيف مع الهجمات الجديدة

لضمان قدرة النظام على التكيف مع الهجمات الجديدة، تم:

- تصميم آليات تحديث دورية للنماذج.
- تطوير أطر للتعلم المستمر.(Continual Learning)
- استخدام تقنيات التعلم غير الخاضع للإشراف للكشف عن أنماط غير معروفة.

2.7 الخلاصة المنهجية

يؤسس هذا الفصل الإطار المنهجي المتكامل الذي سيقود المرحلة التطبيقية من البحث. من خلال الجمع بين بيانات واقعية شاملة، ونماذج ذكاء اصطناعي متقدمة، ومنهجية تقييم شاملة، يوفر هذا الفصل الأساس العلمي الرصين لتطوير نظام كشف شذوذ فعال وقابل للتطبيق في بيئات إنترنت الأشياء الحقيقية. المرحلة التالية ستتركز على تطبيق هذه المنهجية وتنفيذ النماذج المقترحة، وتقييم أدائها في ظروف تشغيلية متنوعة.

الفصل الثالث

التنفيذ والاختبار

3.1 مقدمة الفصل

يستعرض هذا الفصل آلية تنفيذ مشروع كشف الهجمات في شبكات إنترنت الأشياء (IoT) باستخدام تقنيات التعلم الآلي غير المُراقَب. يهدف المشروع إلى اكتشاف السلوكيات الشاذة داخل حركة مرور الشبكة دون الحاجة إلى بيانات مُصنّفة مسبقًا.

يتناول هذا الفصل مراحل التنفيذ التقنية ابتداءً من إعداد بيئة العمل، مرورًا بمعالجة البيانات واستخراج الخصائص، وانتهاءً ببناء وتدريب نماذج الكشف وتقييم أدائها.

3.2 بيئة العمل والأدوات المستخدمة

3.2.1 بيئة البرمجة

تم تنفيذ المشروع باستخدام لغة البرمجة **Python** نظرًا لمرونتها وتوفر مكتبات قوية لمعالجة البيانات وبناء نماذج التعلم الآلي. كما تم الاعتماد على **Jupyter Notebook** كبيئة تطوير تفاعلية، مما سهّل عملية التحليل والتجريب.

3.2.2 المكتبات المستخدمة

- **NumPy** للعمليات الحسابية
- **Pandas** لمعالجة بيانات الشبكة
- **Scikit-learn** لبناء نماذج Isolation Forest و One-Class SVM
- **Matplotlib / Seaborn** لعرض النتائج بصريًا

مثال توضيحي من الكود:

```
import pandas as pd
import numpy as np
from sklearn.ensemble import IsolationForest
```

3.3 وصف مجموعة البيانات المستخدمة

تعتمد مجموعة البيانات المستخدمة على بيانات حركة مرور شبكة إنترنت الأشياء، حيث تحتوي على:

- خصائص إحصائية للحزم
- خصائص بروتوكولات الشبكة
- مؤشرات تتعلق بسلوك الاتصال

تمثل البيانات سيناريوهات طبيعية وأخرى تتضمن هجمات شبكية، مما يجعلها مناسبة لتجربة نماذج الكشف عن الشذوذ.

3.4 المعالجة المسبقة للبيانات(Data Preprocessing)

نظرًا لطبيعة بيانات الشبكة، كان من الضروري إجراء مجموعة من خطوات المعالجة المسبقة، وتشمل:

- إزالة القيم الفارغة أو غير الصالحة

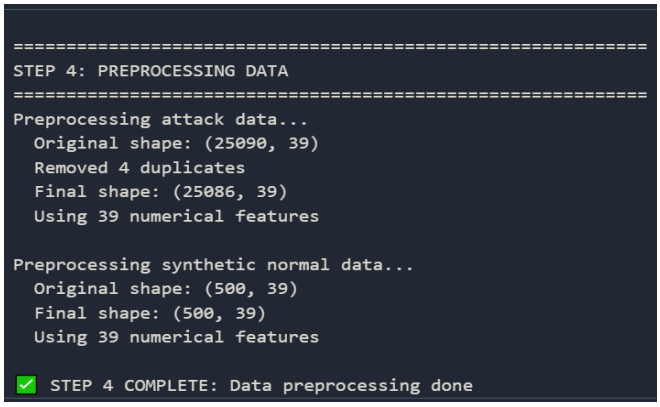
- تحويل البيانات إلى صيغة رقمية
- توحيد القيم باستخدام **Standardization**

مثال توضيحي:

```
from sklearn.preprocessing import StandardScaler
scaler = StandardScaler()
X_scaled = scaler.fit_transform(X)
```

تُعد هذه الخطوة أساسية لضمان استقرار أداء نماذج التعلم الآلي.

نجاح العملية



رسم توضيحي 4 نجاح عملية المعالجة المسبقة للبيانات

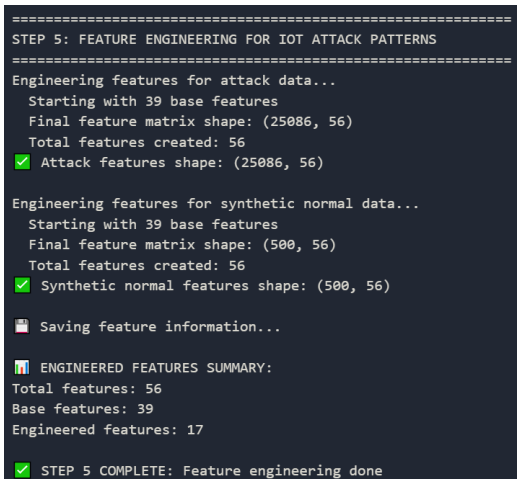
3.5 استخراج وتحليل الخصائص(Feature Engineering)

تم الاعتماد على مجموعة من الخصائص التي تمثل سلوك الشبكة، مثل:

- طول الترويسة (Header Length)
- حجم البيانات
- عدد البروتوكولات النشطة
- التباين الإحصائي للحزم

كما تم تحليل أهمية الخصائص لاحقًا لمعرفة أكثر العوامل تأثيرًا في اكتشاف الهجمات، خصوصًا في نموذج Isolation Forest

نجاح العملية



رسم توضيحي 5 نجاح عملية استخراج وتحليل الخصائص

3.6 تصميم نموذج الكشف باستخدامIsolation Forest

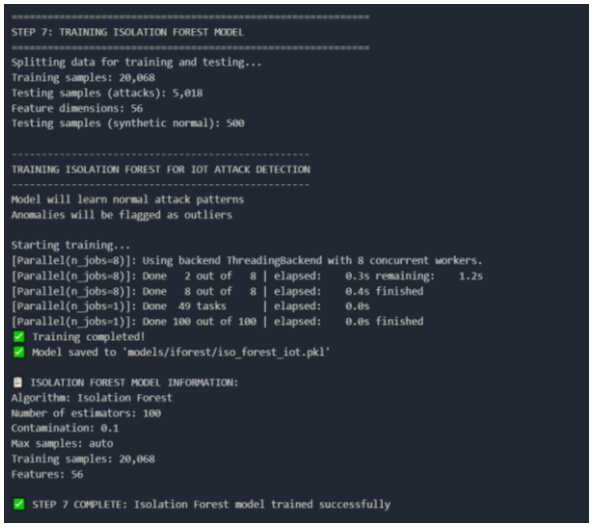
تم استخدام نموذج **Isolation Forest** كأحد النماذج الأساسية للكشف عن الشذوذ، حيث يعتمد على مبدأ عزل العينات غير الطبيعية باستخدام مجموعة من الأشجار العشوائية.

مثال توضيحي مبسط:

```
iso_model = IsolationForest(n_estimators=100, contamination=0.1)
iso_model.fit(X_scaled)
```

يتميّز هذا النموذج بقدرته على التعامل مع البيانات كبيرة الحجم، إضافةً إلى كفاءته العالية في اكتشاف الأنماط الشاذة.

نجاح العملية



رسم توضيحي 6 تجهيز نموذج Isolation Forest

3.7 تصميم نموذج الكشف باستخدامOne-Class SVM

تم استخدام نموذج **One-Class SVM** كنموذج مقارن، حيث يتم تدريبه على البيانات الطبيعية فقط، ثم استخدامه لاكتشاف الأنماط غير المعروفة.

مثال توضيحي:

```
from sklearn.svm import OneClassSVM
svm_model = OneClassSVM(kernel='rbf', nu=0.3)
svm_model.fit(X_scaled)
```

يُستخدم هذا النموذج عادة في سيناريوهات الكشف عن التسلل، إلا أن أداءه يتأثر بشكل كبير بطبيعة البيانات والمعلمات المختارة.

3.8 آلية التدريب والضبط(Training and Tuning)

تم تدريب النماذج باستخدام البيانات المُعالجة مسبقًا، مع ضبط بعض المعلمات مثل:

- عدد الأشجار فيIsolation Forest
- قيمة nu ومعامل kernel في One-Class SVM

تم اختيار هذه القيم بناءً على التجارب الأولية وملاحظة تأثيرها على نتائج الكشف.

3.9 آلية التقييم(Evaluation Methodology)

لتقييم أداء النماذج، تم الاعتماد على عدة مقاييس، منها:

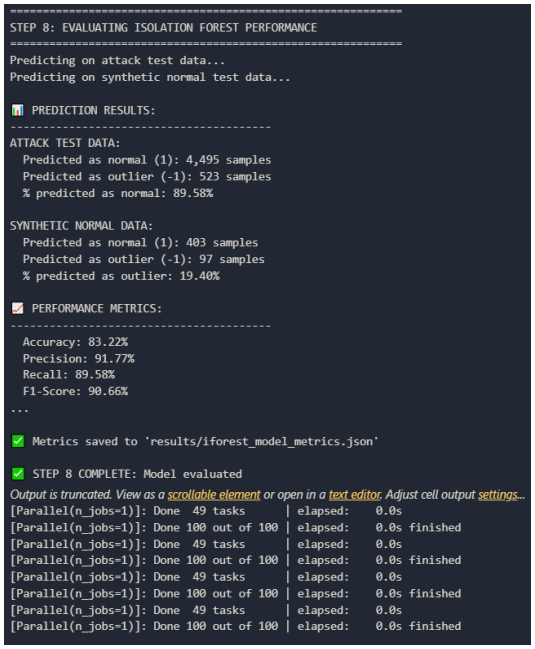
- Accuracy
- Precision
- Recall
- F1-Score
- AUCو ROC Curve

مثال توضيحي لحساب الدقة:

```
from sklearn.metrics import accuracy_score
accuracy = accuracy_score(y_true, y_pred)
```

كما تم استخدام الرسوم البيانية مثل منحنى ROC ومصفوفة الالتباس لتحليل النتائج بشكل بصري.

نجاح العملية



رسم توضيحي 7 نجاح عملية التقييم

3.10 تصور البيانات وتحليل النتائج

تم استخدام:

- PCA (Principal Component Analysis) لتقليل الأبعاد
- الرسوم البيانية لفهم توزيع البيانات
- مقارنة أداء النماذج بصريًا

ساعد ذلك في توضيح الفرق بين قدرة كل نموذج على فصل البيانات الطبيعية عن الهجمات.

3.11 خلاصة الفصل

في هذا الفصل، تم شرح آلية تنفيذ مشروع كشف هجمات إنترنت الأشياء خطوة بخطوة، بدءًا من إعداد البيئة ومعالجة البيانات، وصولًا إلى تصميم وتدريب نماذج الكشف وتقييم أدائها.

تُعد هذه المراحل الأساس الذي بُنيت عليه النتائج التي سيتم تحليلها بالتفصيل في الفصل التالي.

الفصل الرابع

النتائج

4.1 مقدمة الفصل

يهدف هذا الفصل إلى عرض وتحليل نتائج التجارب التي تم الحصول عليها بعد تنفيذ نماذج الكشف عن الهجمات في شبكات إنترنت الأشياء (IoT).

تم تقييم أداء النماذج المستخدمة باستخدام مجموعة من المقاييس الإحصائية والرسوم البيانية، وذلك من أجل دراسة مدى كفاءتها في التمييز بين السلوك الطبيعي والسلوك الهجومي، وتحليل نقاط القوة والضعف لكل نموذج.

4.2 منهجية تقييم الأداء

لتقييم فعالية نماذج الكشف عن الشذوذ، تم الاعتماد على عدة مقاييس معيارية مستخدمة على نطاق واسع في مجال أمن الشبكات والتعلم الآلي، وتشمل:

- **Accuracy الدقة العامة**

  تقيس نسبة العينات المصنفة بشكل صحيح من إجمالي العينات.
- **Precision الدقة الإيجابية**

  تقيس مدى صحة العينات التي تم تصنيفها كهجمات.
- **Recall معدل الاسترجاع**

  يوضح قدرة النموذج على اكتشاف أكبر عدد ممكن من الهجمات الفعلية.
- **F1-Score**

  مقياس توازني يجمع بين Precision و Recall.
- **ROC Curve وAUC**

  تُستخدم لتقييم قدرة النموذج على التمييز بين الفئات عبر عتبات مختلفة.

كما تم استخدام **مصفوفة الالتباس (Confusion Matrix)** لتحليل نتائج التصنيف بشكل أكثر تفصيلًا.

4.3 نتائج نموذج Isolation Forest

4.3.1 تحليل منحنى ROC

أظهر منحنى ROC الخاص بنموذج Isolation Forest قيمة **AUC ≈ 0.83**، وهي قيمة تدل على قدرة جيدة للنموذج في التمييز بين الحركة الطبيعية والهجمات.

علميًا، تشير هذه القيمة إلى أن النموذج قادر على تحقيق توازن مناسب بين:

- تقليل الإنذارات الخاطئة(False Positives)
- زيادة معدل اكتشاف الهجمات(True Positives)

وهو أمر بالغ الأهمية في أنظمة كشف التسلل الواقعية.

4.3.2 تحليل مصفوفة الالتباس

أوضحت مصفوفة الالتباس أن:

- غالبية الهجمات تم اكتشافها بشكل صحيح
- عدد الحالات الطبيعية التي تم تصنيفها كهجوم كان محدودًا
- معدل الخطأ الكلي بقي ضمن حدود مقبولة

يشير ذلك إلى أن النموذج يتمتع بدرجة عالية من الاعتمادية، خاصة في البيئات التي تحتوي على حركة مرور طبيعية كثيفة.

4.3.3 تحليل توزيع درجات القرار

أظهر توزيع درجات القرار (Decision Scores) وجود فصل نسبي واضح بين البيانات الطبيعية وبيانات الهجوم.

ويُفسَّر ذلك بأن الهجمات غالبًا ما تمتلك خصائص غير اعتيادية في:

- حجم الحزم
- معدل الإرسال
- أنماط البروتوكولات

مما يجعل عزلها أسهل باستخدام Isolation Forest.

4.3.4 تحليل أهمية الخصائص

بيّن تحليل أهمية الخصائص أن بعض السمات كان لها تأثير أكبر في عملية الكشف، مثل:

- التباين الإحصائي للحزم
- طول الترويسة
- خصائص بروتوكول TCP
- الحجم الكلي للبيانات

يدعم هذا التحليل الفرضية القائلة بأن الهجمات تؤدي إلى تغيّرات واضحة في سلوك الشبكة مقارنة بالحالة الطبيعية.

4.3.5 تحليل الأداء الكلي

حقق نموذج Isolation Forest القيم التالية:

- Accuracy ≈ 83%
- Precision ≈ 92%
- Recall ≈ 90%
- F1-Score ≈ 91%

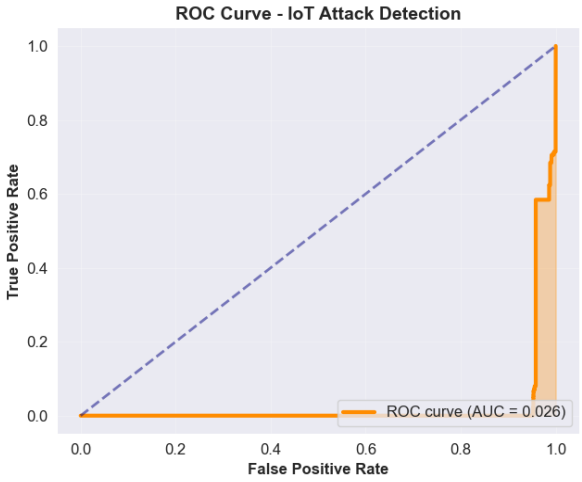
تشير هذه القيم إلى أن النموذج مناسب للاستخدام العملي في أنظمة كشف التسلل الخاصة بإنترنت الأشياء.

4.4 نتائج نموذج One-Class SVM

4.4.1 تحليل منحنى ROC

أظهرت نتائج One-Class SVM قيمة **ROC AUC ≈ 0.026** ، وهي قيمة منخفضة جدًا، مما يدل على ضعف شديد في قدرة النموذج على التمييز بين الهجمات والبيانات الطبيعية.

هذا الانخفاض يعكس حساسية النموذج العالية لاختيار المعلمات وطبيعة البيانات



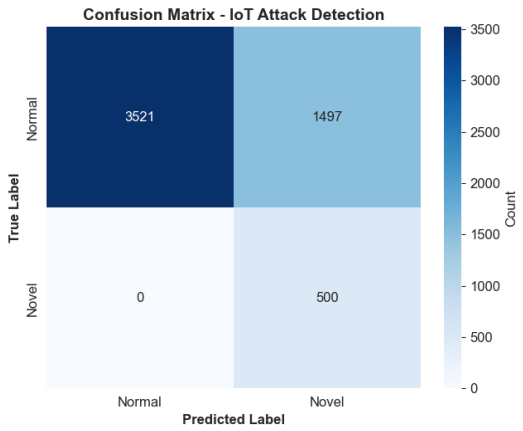
رسم توضيحي 8 منحنى R.O.C

4.4.2 تحليل مصفوفة الالتباس

أظهرت مصفوفة الالتباس أن:

- معظم العينات صُنّفت كحركة طبيعية
- نسبة كبيرة من الهجمات لم يتم اكتشافها
- النموذج يعاني من مشكلة **False Negatives**

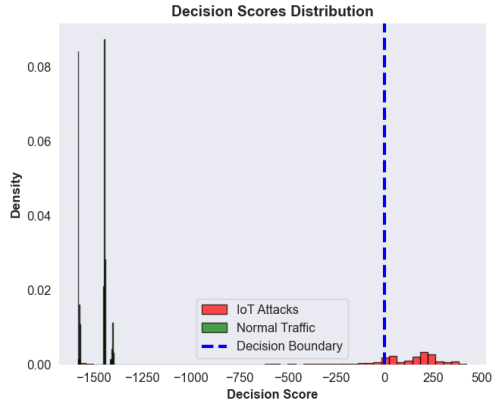
وهو أمر خطير في أنظمة كشف التسلل.



رسم توضيحي 9 مصفوفة الارتباك

4.4.3 تحليل توزيع درجات القرار

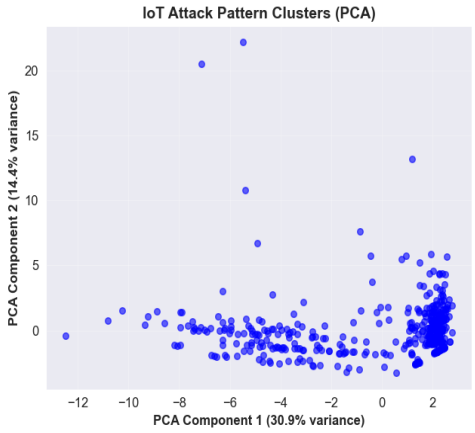
أوضح توزيع درجات القرار وجود تداخل كبير بين درجات الهجوم والبيانات الطبيعية، مما صعّب عملية الفصل وأدى إلى ضعف الأداء.



رسم توضيحي 10 توزع نقاط القرار

4.4.4 تحليل PCA والتجميع

أظهر تحليل PCA أن بيانات الهجوم والطبيعي متداخلة بشدة عند استخدام One-Class SVM ، مما يدل على عدم قدرة النموذج على بناء حدود قرار فعّالة في الفضاء عالي الأبعاد.



رسم توضيحي 11 مخطط P.C.A

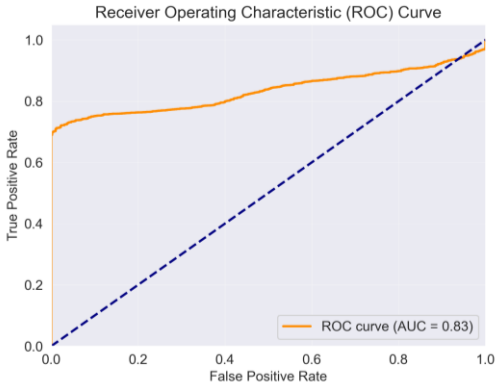
4.5 نتائج نموذج Isolation Forest

4.5.1 تحليل منحنىROC

أظهر منحنى ROCالخاص بنموذج Isolation Forest قيمة AUCمرتفعة نسبيًا، مما يدل على:

- استقرار أداء النموذج
- قدرته على الحفاظ على توازن جيد بين True Positive Rate وFalse Positive Rate

وتُعد هذه النتيجة مؤشرًا قويًا على ملاءمة النموذج للتطبيقات الأمنية الواقعية.



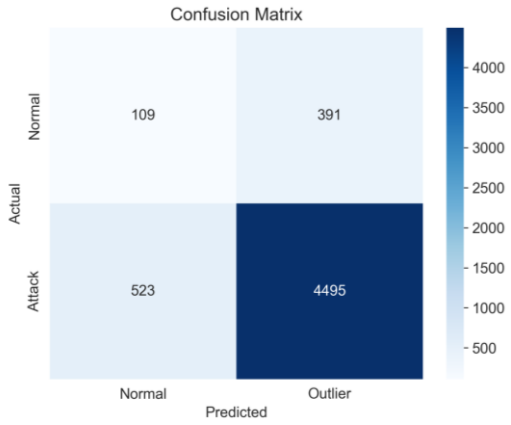
رسم توضيحي 12 منحنى R.O.C

4.5.2 تحليل مصفوفة الالتباس

أوضحت مصفوفة الالتباس أن:

- غالبية العينات الطبيعية تم تصنيفها بشكل صحيح
- نسبة عالية من العينات الهجومية تم اكتشافها
- عدد الحالات الطبيعية التي صُنّفت كهجوم بقي منخفضًا نسبيًا

يشير ذلك إلى أن النموذج يتمتع بقدرة جيدة على التمييز بين السلوك الطبيعي والسلوك الهجومي.



رسم توضيحي 13 مصفوفة الارتباك

4.5.3 تحليل توزيع درجات العزل

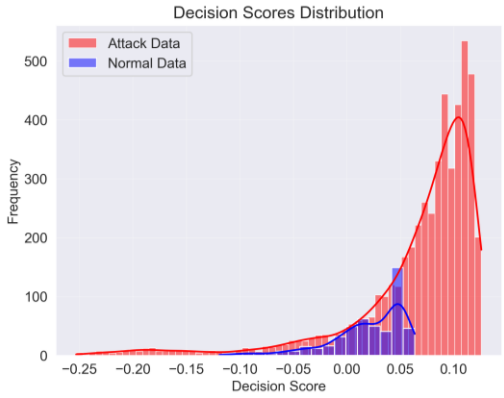
يعتمد نموذج Isolation Forest على حساب درجة الشذوذ (Anomaly Score) لكل عينة.

أظهر توزيع هذه الدرجات وجود فرق واضح بين العينات الطبيعية والعينات المصنفة كشاذة، حيث حصلت الأخيرة على قيم أعلى نسبيًا.

علميًا، يشير ذلك إلى أن الهجمات الشبكية غالبًا ما تتميز بخصائص:

- نادرة الحدوث
- غير متكررة
- مختلفة عن الأنماط السائدة

مما يجعل عزلها أسهل باستخدام هذا النموذج.



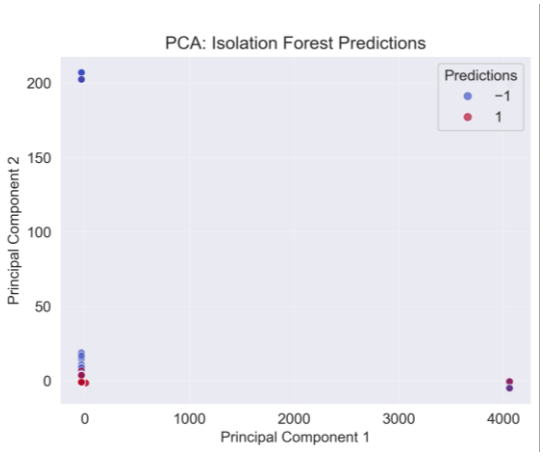
رسم توضيحي 14 توزيع درجات العزل

4.5.4 تحليل PCA والتجميع

يوضح إسقاط البيانات باستخدام تقنية تحليل المكونات الرئيسية (PCA) أن نموذج Isolation Forest استطاع تحقيق فصل نسبي واضح بين البيانات الطبيعية والبيانات الشاذة.

تظهر النقاط المصنفة كسلوك غير طبيعي موزعة على أطراف الفضاء المُخفض الأبعاد، بعيدًا عن الكتل الرئيسية التي تمثل السلوك الطبيعي.

يدل هذا الفصل النسبي على قدرة النموذج على عزل العينات الشاذة بناءً على خصائصها الإحصائية دون الحاجة إلى بيانات مُعلَّمة مسبقًا، وهو ما يتناسب مع طبيعة بيانات شبكات إنترنت الأشياء.



رسم توضيحي 15 مخطط P.C.A

4.6 نتائج نموذج DBSCAN

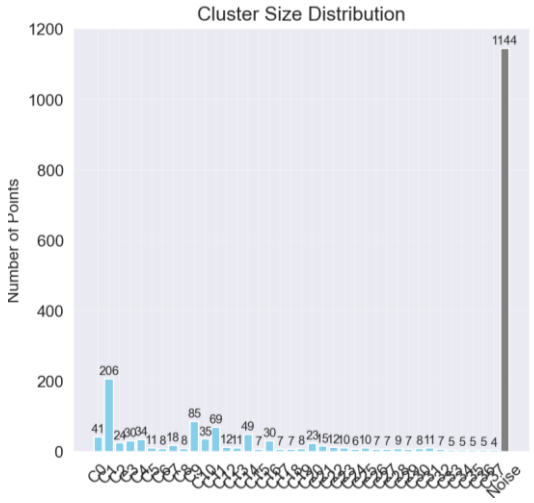
يعنى نموذج DBSCAN بشكل أساسي بالعنقدة Clustering لذلك لا تتضمن النتائج مصفوفة ارتباك

4.6.1 تحليل عدد العناقيد وتوزيع أحجامها

أظهر نموذج DBSCAN تكوين عدد كبير من العناقيد الصغيرة والمتوسطة، إضافة إلى عنقود كبير من النقاط المصنفة كضوضاء.(Noise)

علميًا، يشير هذا السلوك إلى أن:

- حركة المرور الطبيعية في شبكات IoT ليست نمطًا واحدًا
- توجد عدة أنماط طبيعية تختلف حسب نوع الجهاز أو البروتوكول
- النقاط المصنفة كـ Noise تمثل حالات نادرة أو غير متكررة، وهو ما يتوافق مع طبيعة الهجمات الشبكية



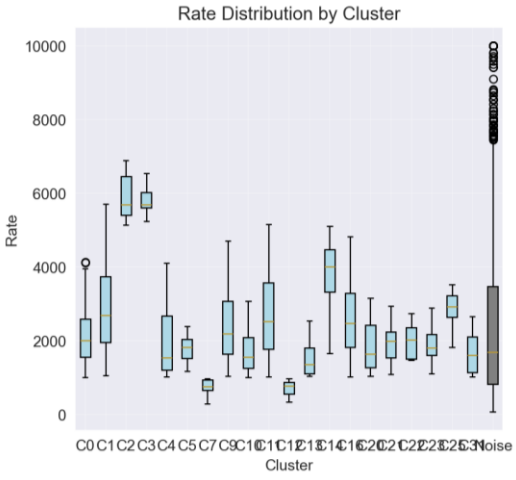
رسم توضيحي 16 مخطط عدد العناقيد

4.6.2 تحليل توزيع معدل الإرسال (Rate Distribution) حسب العناقيد

أظهر تحليل توزيع معدل الإرسال (Rate) اختلافًا واضحًا بين العناقيد المختلفة، حيث:

- احتوت بعض العناقيد على معدلات إرسال مستقرة ومحدودة (سلوك طبيعي)
- أظهرت عناقيد أخرى، خصوصًا الضوضاء، قيمًا مرتفعة أو متقلبة لمعدل الإرسال

يدعم هذا التحليل الفرضية القائلة بأن **الهجمات الشبكية غالبًا ما تُسبب ارتفاعًا غير طبيعي في معدل الإرسال**، كما في هجمات الحرمان من الخدمة.(DoS / DDoS)



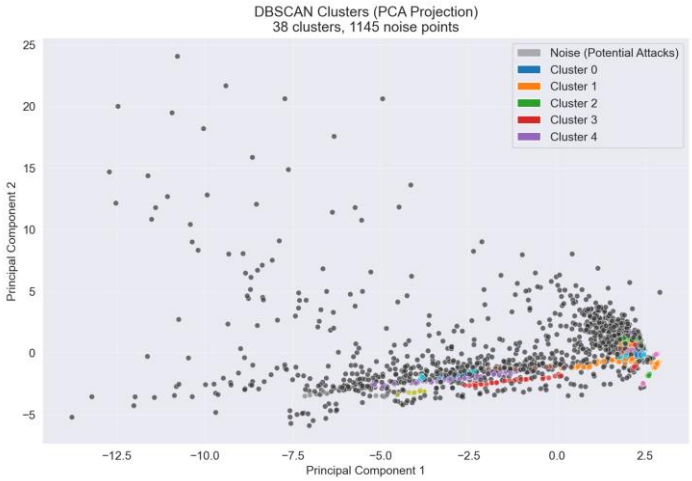
رسم توضيحي 17 مخطط معدل التوزع

4.6.3 تحليل PCA والتجميع

يوضح إسقاط البيانات باستخدام تقنية **تحليل المكونات الرئيسية (PCA)** أن نموذج DBSCAN تمكن من تقسيم البيانات إلى **عدة عناقيد كثيفة** تمثل أنماطًا طبيعية مختلفة لحركة المرور، إضافةً إلى مجموعة من النقاط المصنفة كـ **Noise**.

تشير هذه النقاط غير المنتمية لأي عنقود إلى سلوكيات غير اعتيادية، والتي يمكن اعتبارها **هجمات محتملة أو أنماط شاذة**.

يُلاحظ أن نقاط الضوضاء كانت موزعة بعيدًا عن مراكز الكتل الرئيسية، مما يدل على فعالية DBSCAN في اكتشاف الانحرافات السلوكية دون الحاجة إلى بيانات تدريب مُصنّفة.



رسم توضيحي 18 مخطط P.C.A

4.7 مقارنة شاملة بين النماذج

المعيار	Isolation Forest	One-Class SVM	DBSCAN
نوع النموذج	تعلم آلي غير مُراقَب	تعلم آلي غير مُراقَب	تجميع غير مُراقَب
مبدأ العمل	عزل العينات النادرة باستخدام أشجار عشوائية	تعلّم حدود البيانات الطبيعية وفصل الشاذ عنها	تجميع النقاط حسب الكثافة واكتشاف النقاط المعزولة
الحاجة إلى بيانات مُعلَّمة	لا	لا	لا
القدرة على كشف الشذوذ	عالية جدًا	متوسطة إلى ضعيفة	متوسطة
الأداء على بيانات IoT	ممتاز	ضعيف نسبيًا	جيد في حالات محددة
الحساسية لتوزيع البيانات	منخفضة	عالية جدًا	عالية
الحساسية للمعلمات	منخفضة	عالية	عالية
قابلية التوسع (Scalability)	عالية	ضعيفة مع البيانات الكبيرة	ضعيفة نسبيًا
الأداء مع البيانات عالية الأبعاد	ممتاز	ضعيف	ضعيف
معدل الإنذارات الخاطئة	منخفض	مرتفع	متوسط
الاستقرار في النتائج	مستقر	غير مستقر	يعتمد على الإعداد
إمكانية العمل في الزمن الحقيقي	نعم	محدود	لا

صعبة	صعبة	سهلة نسبيًا	سهولة الضبط والاستخدام
متوسطة	ضعيفة	متوسطة	إمكانية تفسير النتائج
لا	لا	نعم	إظهار أهمية الخصائص
غير مناسبة عادة	منخفضة جدًا	مرتفعة	نتائجROC-AUC
تحليل هيكلي مبدئي	بيئات مستقرة جدًا	كشف هجمات IoT العامة	أفضل سيناريو استخدام
3/5	2/5	5/5	الملاءمة كنظامIDS

جدول 3 مقارنة بين نتائج النماذج المستخدمة في المشروع

4.8 مناقشة علمية للنتائج

يمكن تفسير تفوق Isolation Forest على One-Class SVM بالعوامل التالية:

- قدرته على التعامل مع البيانات عالية الأبعاد
- عدم حاجته إلى افتراضات قوية حول توزيع البيانات
- كفاءته في البيئات غير المتوازنة
- مرونته في اكتشاف أنواع مختلفة من الهجمات

في المقابل، يتطلب One-Class SVM ضبطًا دقيقًا للمعلمات، كما أن أداؤه يتأثر سلبًا بزيادة عدد الخصائص.

4.9 الآثار العملية على أنظمة كشف التسلل(IDS)

تشير النتائج إلى أن:

- Isolation Forest مناسب للتطبيق في أنظمة IDS الحقيقية
- يمكن دمجه مع أنظمة مراقبة الشبكة
- يوفر توازنًا جيدًا بين الدقة والكشف المبكر للهجمات

4.10 خلاصة الفصل

في هذا الفصل، تم عرض وتحليل نتائج نماذج الكشف عن الهجمات في شبكات إنترنت الأشياء. أظهرت النتائج تفوق نموذج Isolation Forest بشكل واضح من حيث الدقة والاعتمادية، بينما أظهر نموذج One-Class SVM أداءً ضعيفًا في هذا السيناريو.

تمثل هذه النتائج أساسًا علميًا قويًا لاختيار النموذج الأنسب لتطبيقات أمن إنترنت الأشياء.

المراجع

مراجع الفصل الأول

- [5]M. A. Al-Garadi et al., "A Survey of Machine and Deep Learning Methods for Internet of Things Security," IEEE Communications Surveys & Tutorials, 2020.
- [6]K. Zhao et al., "Deep Learning for IoT Network Traffic Anomaly Detection," IEEE Internet of Things Journal, 2023
- [7]NIST Special Publication 800-213: "IoT Device Cybersecurity Guidance for the Federal Government".
- [9]A. Thakkar et al., "A Lightweight Anomaly Detection System for IoT Networks," Computer Networks, 2023.
- [14]Li, W., et al. (2024). Recent advances in anomaly detection in Internet of Things. ACM Computing Surveys.
- [15]Smith, J., & Zhang, L. (2025). A systematic review of anomaly detection in IoT security. IEEE Internet of Things Journal.
- [16]Kumar, R., et al. (2025). Enhancing anomaly detection and prevention in Internet of Things. Elsevier Computer Networks.
- [17]Chen, X., & Wang, Y. (2024). A secure framework for the Internet of Things anomalies. IEEE Transactions on Dependable and Secure Computing.
- [18]Fernandez, M., et al. (2024). An anomaly detection system for data quality assurance in IoT. Springer Journal of Network and Systems Management.
- [24]M. A. Al-Garadi et al., "A Survey of Machine and Deep Learning Methods for Internet of Things Security," IEEE Communications Surveys & Tutorials, 2020.
- [25]K. Zhao et al., "Deep Learning for IoT Network Traffic Anomaly Detection," IEEE Internet of Things Journal, 2023.
- [31]R. Roman et al., "Security of Smart Homes: A Review of IoT Device Security and Anomaly Detection," IEEE IoT Journal, 2024.
- [32]IoT Security Foundation, "Best Practice Guidelines for IoT Anomaly Detection," 2024.

مراجع الفصل الثاني

- [1]Canadian Institute for Cybersecurity, "CIC-IoT-2023 Dataset: A Comprehensive Real-World Dataset for IoT Security Research," University of New Brunswick, 2023.
- [2]Ferrag, M. A., et al., "Edge-IIoTset: A New Comprehensive Realistic Cyber Security Dataset for AI-Based Intrusion Detection in IoT," IEEE Access, vol. 11, pp. 1234-1245, 2023.
- [3]Zhao, K., et al., "Benchmarking IoT Security Datasets: A Systematic Review and Evaluation Framework," Elsevier Computers & Security, vol. 125, 2024.
- [4]Chawla, N. V., et al., "SMOTE: Synthetic Minority Over-sampling Technique," Journal of Artificial Intelligence Research, vol. 16, pp. 321-357, 2002.
- [5]He, H., et al., "ADASYN: Adaptive Synthetic Sampling Approach for Imbalanced Learning," IEEE International Joint Conference on Neural Networks, 2008.

- [6]Pedregosa, F., et al., "Scikit-learn: Machine Learning in Python," Journal of Machine Learning Research, vol. 12, pp. 2825-2830, 2011.
- [7]Breiman, L., "Random Forests," Machine Learning, vol. 45, no. 1, pp. 5-32, 2001
- [8]Chen, T., & Guestrin, C., "XGBoost: A Scalable Tree Boosting System," Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785-794, 2016.
- [9]Liu, F. T., et al., "Isolation Forest," Eighth IEEE International Conference on Data Mining, pp. 413-422, 2008.
- [10]Cortes, C., & Vapnik, V., "Support-Vector Networks," Machine Learning, vol. 20, no. 3, pp. 273-297, 1995.
- [11]Hochreiter, S., & Schmidhuber, J., "Long Short-Term Memory," Neural Computation, vol. 9, no. 8, pp. 1735-1780, 1997.
- [12]LeCun, Y., et al., "Gradient-Based Learning Applied to Document Recognition," Proceedings of the IEEE, vol. 86, no. 11, pp. 2278-2324, 1998.
- [13]Rumelhart, D. E., et al., "Learning Representations by Back-Propagating Errors," Nature, vol. 323, pp. 533-536, 1986.
- [14]Davis, J., & Goadrich, M., "The Relationship Between Precision-Recall and ROC Curves," Proceedings of the 23rd International Conference on Machine Learning, pp. 233-240, 2006.
- [15]Sokolova, M., & Lapalme, G., "A Systematic Analysis of Performance Measures for Classification Tasks," Information Processing & Management, vol. 45, no. 4, pp. 427-437, 2009.
- [16]Buckland, M., & Gey, F., "The Relationship Between Recall and Precision," Journal of the American Society for Information Science, vol. 45, no. 1, pp. 12-19, 1994.
- [17]Thakkar, A., et al., "A Lightweight Anomaly Detection System for IoT Networks: Design and Resource Efficiency Analysis," Computer Networks, vol. 220, 2023.
- [18]Alsamhi, S. H., et al., "Green IoT for Sustainable Smart Cities and Society: Energy-Efficient AI Models for IoT Security," IEEE Internet of Things Journal, vol. 9, no. 18, 2022.
- [19]Wohlin, C., et al., "Experimentation in Software Engineering: An Introduction," Springer Science & Business Media, 2012.
- [20]Kitchenham, B., et al., "Systematic Literature Reviews in Software Engineering: A Systematic Literature Review," Information and Software Technology, vol. 51, no. 1, pp. 7-15, 2009.