مجلـــة كليـــة التربيـــة طرابلـــس

Journal of the Faculty of Education Tripoli

نحو ذكاء اصطناعي مفسَّر لكشف الشذوذ في الشبكات الاجتماعية: دراسة مقارنة لنموذج شجرة القرار في بيئة فيسبوك

مروى بهجات سويدان[1]، رمزي حميد القانوني [2]، عبدالناصر عبدالحميد أحمد ضياف[2]

[1] قسم الحاسوب ، كلية التربية طرابلس ، جامعة طرابلس، طرابلس ، ليبيا.

[2] قسم تقنيات الانترنت ، كلية تقنية المعلومات ، جامعة طرابلس، طرابلس ، ليبيا.

البريد الإلكتروني للمؤلف المكلف بالتواصل: ma.swidan@uot.edu.ly

Article history

Received: Month Mar, 2026 Accepted: Month Mar, 2026

الملخص:

أصبح العالم بأجمعه منغمسًا في وسائل التواصل الاجتماعي رغم كل التهديدات التي تواجهه، وهذه التهديدات تختلف قوتها والهدف منها من شخص إلى آخر، وقد تندرج الأنماط غير الطبيعية والتي تعرف بالشذوذ او الانماط الشاذة تحت هذه التهديدات. وقد اتجهت العديد من الأبحاث لمحاولة الكشف عن الأنماط الشاذة وغير الشاذة بعدة طرق. تناولت هذه الدراسة تحدي كشف الأنماط الشاذة في الشبكات الاجتماعية الرقمية، باستخدام مجموعة بيانات Facebook Social Circles، اقترحنا إطار عمل يعتمد على تعلم الآلة ويستفيد من مقاييس الرسم البياني الهيكلية (مثل المركزية ومعاملات التجمع) لتحديد الشذوذ. تم إجراء تحليل مقارن شامل بين عشر خوارزميات مختلفة لتعلم الآلة، بما في ذلك اشجار القرار، والشبكات العصبية، وتعزيز التدرج (Gradient Boosting). أظهرت النتائج التجريبية أن نموذج شجرة القرار (Decision Tree) حقق دقة متميزة بلغت 98.50% مع مقياس F1 قدره 0.98. وبالإضافة إلى أدائه العالي، يوفر النموذج المقترح قابلية عالية للتفسير، مما يتيح فهماً واضحاً للمنطق الكامن وراء كشف الشذوذ. يخلص هذا البحث إلى أن الجمع بين الميزات القائمة على الرسم البياني ونماذج الذكاء الاصطناعي القابلة للتفسير يوفر حلاً قوياً وموثوقاً لتأمين بيئات وسائل التواصل الاجتماعي من الانماط الشاذة.

الكلمات المفتاحية: كشف الشذوذ، الانماط الشاذة، تعلم الآلة، أشجار القرار، الشبكات الاجتماعية، الذكاء الاصطناعي، مقاييس الرسم البياني.

Towards Explainable Artificial Intelligence for Anomaly Detection in Social Networks: A Comparative Study of Decision Tree Models in Facebook Environment

A B S T R A C T:

Today, the world is deeply immersed in social media platforms despite the escalating threats they pose. These threats vary in intensity and objectives, with abnormal patterns—commonly referred to as "anomalies"—falling under these significant security risks. Numerous studies have explored various methods to detect anomalous and non-anomalous patterns. This study addresses the challenge of anomaly detection in digital social networks using the Facebook Social Circles dataset. We propose a machine learning framework that leverages structural graph metrics, such as centrality and clustering coefficients, to identify anomalies. A comprehensive comparative analysis was conducted among ten different machine learning algorithms, including Decision Trees, Neural Networks, and Gradient Boosting. Experimental results demonstrated that the Decision Tree model achieved an outstanding accuracy of 98.50% with an F1-score of 0.98. In addition to its high performance, the proposed model is highly explainable, providing a clear understanding of the underlying logic behind anomaly detection. This research concludes that combining graph-based features with Explainable AI (XAI) models provides a robust and reliable solution for securing social media environments against anomalous patterns.

Keywords: Anomaly Detection, Anomalous Patterns, Machine Learning, Decision Trees, Social Networks, Artificial Intelligence, Graph Metrics.

المقدمة:

تُمثل شبكات التواصل الاجتماعي وعلى رأسها منصة فيسبوك عنصراً حيوياً في الحياة الرقمية المعاصرة، حيث تشكل مساحات للتفاعل الاجتماعي والاقتصادي والسياسي لعدد كبير جدا من الاشخاص، هذا التوسع الهائل رافقه ارتفاع متزايد في التهديدات الأمنية التي تستغل طبيعة المنصات المفتوحة التي ينشر من خلالها الاشخاص اخبارهم و افكارهم وتوجهاتهم واحاسيسهم وغيرها من المعلومات. وهناك خيط رفيع للتمييز ما اذا كان هذا السلوك شاذ او غير شاذ فعلى سبيل المثال النشر على الفيسبوك هو نمط غير شاذ بطبيعته ولكن النشر المتسارع في فترة قصيرة جدا قد يعتبر نمط شاذ، وعلى نفس السياق فان طلبات الصداقة هي نمط غير شاذ ولكن عندما تكون بشكل كبير وفي فترة وجيزة فهذا الامر شاذ. إن الفشل في التمييز بين االنمط الطبيعي والأنماط الشاذة يؤدي إلى تزعزع الثقة وانتشار المعلومات المضللة، وخسائر مالية للشركات والأفراد على حد سواء[1]،[9]. إن كشف الشذوذ في الشبكات الاجتماعية لا يقتصر على رصد السلوكيات الفردية، بل يمتد ليشمل تحليل 'الانحراف الهيكلي' عن المجتمعات الرقمية المستقرة، وهو ما أكدته دراسة (Zardi) في إطارها الموحد لكشف المجتمعات والشذوذ [15] .

في هذا السياق، أصبح التنبؤ السريع والوصول الفعّال لهذه الأنماط وتصنيفها بدقة أمراً بالغ الأهمية حيث ان هذه الأنماط الشاذة تتطور باستمرار لتتحايل على آليات الكشف التقليدية، بالإضافة الى أن التمييز بين السلوك "الشاذ" و "الطبيعي" يعتبر تحدي كبير في بيئة ديناميكية ومتغيرة مثل فيسبوك على الرغم من أن نماذج التعلم العميق قد أظهرت قدرات عالية في التصنيف، إلا أنها غالباً ما تفتقر إلى الشفافية وقابلية التفسير (Interpretability)، مما يجعل من الصعب على محللي الأمن فهم الأساس المنطقي وراء قرارات الكشف [3]،[10].

تعتمد المنهجيات التقليدية في كشف الشذوذ غالباً على تحليل البيانات النصية أو السلوكية المباشرة، إلا أن الدراسات الحديثة، ومنها دراسة، (Helmi, R. A. A. et al., 2022) أثبتت أن مقاييس الرسم البيانيGraph Metrics تمثل أداة أكثر عمقاً وفاعلية. فمن خلال تمثيل الشبكة الاجتماعية كرسم بياني (Graph) يتكون من عقد (Nodes) تمثل المستخدمين وروابط (Edges) تمثل التفاعلات، يمكننا استخراج ميزات هيكلية مثل درجة المركزية (Degree Centrality) ومعامل التكتل (Clustering Coefficient) التي تكشف عن جوهر السلوك البشري مقارنة بالسلوك الآلي أو الشاذ[5]. ومع ذلك، تواجه معظم نماذج التعلم الآلي المتقدمة، مثل الشبكات العصبية(Neural Networks) وغيرها مشكلة الصندوق الأسود (Black Box) حيث تعطي نتائج دقيقة دون توضيح الأسباب الكامنة وراء قرار التصنيف. هنا تبرز فجوة بحثية تتمثل في الحاجة إلى نماذج لا تكتفي بالدقة الرقمية فحسب، بل تقدم تفسيراً منطقياً (Explainability) للقرارات المتخذة.

يأتي هذا البحث ليعالج هذه الفجوة عبر استخدام نموذج شجرة القرار(Decision Tree) حيث ان هذا النموذج له قدرة على محاكاة المنطق البشري من خلال قواعد قرار (If–Then Rules) واضحة ومبنية على

معايير إحصائية دقيقة مثل(Entropy) [10]. يهدف هذا البحث إلى دراسة تأثير مقاييس الشبكة الاجتماعية في كشف الأنماط الشاذة وغير الشاذة، مع التركيز على تقديم نموذج مفسَّر(Explainability) يتفوق في دقته ومنطقيته على النماذج المعقدة الأخرى، مستنداً إلى النتائج التجريبية التي أظهرت كفاءة استثنائية لشجرة القرار في هذا النطاق. على الرغم من دقة النماذج المعقدة مثل الشبكات العصبية في أبحاث السلوك الشاذ السابقة، إلا أنها تظل نماذج غير مفسَّرة(Black-box) ، مما يصعب على محللي الأمن فهم أسباب التصنيف. تبرز فجوة البحث الحالية في الحاجة إلى نموذج يجمع بين الدقة العالية والشفافية المنطقية، وهو ما تسعى هذه الدراسة لتحقيقه من خلال تطوير وتطبيق نموذج مُحسَّن لشجرة القرار لكشف الأنماط الشاذة وغير الشاذة على منصة فيسبوك.

مما سبق نجد ان مساهمة هذا البحث تكمن في تقديم إطار عمل يجمع بين خصائص الشبكة الهيكلية (Graph Metrics) والذكاء الاصطناعي المفسر (XAI)، حيث لا يكتفي النموذج بكشف الشذوذ بدقة عالية، بل يستخرج قواعد قرار (Decision Rules) تمكن مسؤولي الأمن من فهم الدوافع البنيوية وراء تصنيف أي عقدة كشاذة، وهو ما يحل معضلة 'الصندوق الأسود' في الأنظمة الأمنية.

الأدبيات السابقة

ان النمط الشاذ او ما يعرف بالشذوذ في شبكات التواصل الاجتماعي يمكن وصفه على أنه نمط سلوكي يختلف بشكل كبير عن الغالبية العظمى من الأنماط الأخرى التي تُعتبر طبيعية أو غير شاذة .(Non-Anomalous) هذا الشذوذ عادة ما يشير إلى أنشطة ضارة مثل الحسابات المزيفة، ورسائل البريد العشوائي، وهجمات التصيد. وقد شهد العقد الأخير تطوراً ملحوظاً في استخدام تقنيات التعلم الآلي للكشف عن الأنشطة غير الطبيعية في الشبكات الاجتماعية. يمكن تقسيم الدراسات السابقة إلى ثلاثة اتجاهات رئيسية:

1. **كشف الشذوذ باستخدام مقاييس الرسم البياني** :(Graph Metrics Analysis) ركزت العديد من الدراسات على أن الخصائص الهيكلية للشبكة أكثر دقة في كشف الشذوذ من المحتوى النصي. ويعزز هذا التوجه ما أشار إليه (Akoglu et al., 2015) إلى أن استخدام مقاييس المركزية (Centrality) ومعامل التكتل (Clustering Coefficient) يمكن أن يكشف عن البوتات (Bots) والحسابات الوهمية التي تحاول محاكاة السلوك البشري ولكنها تفشل في تقليد البنية الهيكلية للتفاعلات الحقيقية [4]. وقد عززت دراسة (Helmi, R. A. A. et al., 2022) هذا الاتجاه من خلال إثبات أن النماذج القائمة على هذه المقاييس تحقق دقة تفوق الطرق التقليدية [5]. وفي دراسة مسحية حديثة وشاملة أكد (Ho et al., 2025) وفي هذا الإطار، قدمت دراسة (Zardi) نموذجاً منهجياً يعتمد على مفهوم الانحراف (Deviation-based) للكشف الموحد عن المجتمعات والأنماط الشاذة، مؤكدة أن الشذوذ الهيكلي يبرز بوضوح كانحراف عن المعايير التنظيمية للمجموعات الاجتماعية داخل الشبكة. [15] وتبرز أهمية الدراسة الحالية في كونها امتداداً تطبيقياً لهذا المفهوم؛ فبينما قدمت (Zardi) إطاراً نظرياً للانحراف الهيكلي، يقوم بحثنا بتحويل هذه الانحرافات الى قيم كمية ملموسة باستخدام مقاييس الرسم البياني مثل المركزية ومعامل التجمع ومن ثم تغذيتها لنموذج شجرة القرار.

2. نماذج التعلم الآلي وتحدي الصندوق الأسود: (Black Box Challenge) استخدمت دراسات أخرى نماذج معقدة مثل الشبكات العصبية العميقة (Deep Learning) أو المجموعات العشوائية (Random Forests) للحصول على دقة عالية. ومع ذلك، ناقشت دراسة (Lundberg & Lee, 2017) أن هذه النماذج تعاني من ضعف قابلية التفسير، مما يجعل من الصعب على مديري الشبكات فهم سبب تصنيف حساب ما كشاذ. هذا النقص في التفسير يقلل من الثقة في اتخاذ قرارات تلقائية مثل حظر المستخدمين [7]. وهو ما أبرزته ايضا دراسة (Arrieta et al., 2020) كفجوة بحثية كبرى تستوجب التحول نحو نماذج شفافة [12]. وفي نفس السياق قدمت دراسة(Lunawat et al. 2023) نموذج GridBoost وهو مصنف يعتمد على تقنيات التعزيز لتحسين دقة كشف الشذوذ في الشبكات الاجتماعية. وعلى الرغم من النجاح الكبير الذي حققه هذا النموذج في تقليل نسب الخطأ والوصول إلى دقة عالية، إلا أنه يمثل نموذجاً للأنظمة المعقدة (Black Box) التي يصعب على البشر فهم آلية اتخاذ القرار داخلها [11]. وفي توجه تقني متقدم، قدمت دراسة (Babu et al. 2026) نموذجاً يعتمد على الشبكات التلافيفية الرسومية غير المحلية (CNLGCN) لكشف شذوذ العقد، مؤكدة على ضرورة دمج الخصائص الهيكلية العميقة للشبكة؛ وبالرغم من الكفاءة العالية لهذه النماذج العميقة في التعامل مع البيانات الضخمة، إلا أنها تزيد من تعقيد عملية التفسير الأمني [14]. وهذا ما يؤكد الفجوة التي يسعى البحث الحالي لسدها؛ وهي كيفية الوصول إلى مستويات دقة تقترب من النماذج المتقدمة مثل GridBoost و CNLGCN ولكن مع الحفاظ على البساطة والشفافية التي توفرها شجرة القرار.

3. التوجه نحو الذكاء الاصطناعي القابل للتفسير (XAI) وشجرة القرار: للتغلب على تحدي الصندوق الأسود، برز اتجاه حديث لاستخدام نماذج قابلة للتفسير. أظهرت أبحاث (Breiman, 2017) أن شجرة القرار (Decision Tree) توفر موازنة ممتازة بين الدقة وقابلية التفسير، حيث يمكن تحويل مسارات الشجرة إلى قواعد منطقية (If–Then Rules) مفهومة للبشر[6]. وقد عززت دراسة(Suhail et al. 2022) هذا التوجه بتأكيدها أن نماذج شجرة القرار توفر موازنة مثالية بين الدقة العالية وقابلية التفسير المنطقي، مما يتيح لمديري الشبكات الاجتماعية ليس فقط اكتشاف الحسابات الشاذة بل وفهم الأسباب الجذرية الهيكلية والسلوكية لهذا الشذوذ، وهو ما يرفع من موثوقية وثقة القرارات الأمنية المتخذة.

يتميز هذا البحث بدمج هذه الاتجاهات، حيث نستخدم مقاييس الرسم البياني كما في دراسة Elghanuni كمدخلات لنموذج شجرة القرار لتحقيق التفسير، لتقديم حل دقيق ومفهوم لكشف الشذوذ في فيسبوك.

منهجية العمل

تعتمد هذه الدراسة على منهجية تجريبية مقارنة للكشف عن الأنماط الشاذة في الشبكات الاجتماعية، مع التركيز بشكل أساسي على مبدأ القابلية للتفسير (Explainability). وقد تم تنفيذ التجارب باستخدام لغة Python3.9 مع الاعتماد على مكتبة NetworkX لمعالجة الرسوم البيانية واستخراج الخصائص الهيكلية، بالاضافة الى مكتبة Scikit–learn لبناءوتقييم خوارزميات تعلم الآلة.

و تم تنفيذ العمل وفق الخطوات المنهجية الموضحة أدناه:

1. وصف مجموعة البيانات (Dataset Description)

تم الاعتماد في هذه الدراسة على مجموعة بيانات Facebook Social Circles المتاحة عبر مستودع(Stanford Large Network Dataset Collection). تُمثل هذه البيانات شبكات اجتماعية واقعية من منصة فيسبوك، حيث تتكون من عقد (Nodes) تمثل المستخدمين، وروابط (Edges) تمثل علاقات الصداقة والتفاعل.

2. استخراج الميزات الهيكلية (Feature Engineering)

لتحويل هيكل الشبكة الخام إلى بيانات قابلة للتحليل بواسطة خوارزميات تعلم الآلة، تم استخراج مجموعة من مقاييس الرسم البياني (Graph Metrics). تعمل هذه الميزات كمؤشرات سلوكية قادرة على تمييز الأنماط الطبيعية عن الشاذة، وشملت هذه المقاييس:

- المركزية (Centrality Measures).
- معامل التجميع (Clustering Coefficient).
- خصائص الجوار الهيكلي للعقد.

3. المعالجة المسبقة للبيانات (Data Pre-processing)

في هذه المرحلة تمر البيانات بعدة خطوات لضمان جودتها قبل مرحلة التدريب:

- تنظيف البيانات :معالجة القيم المفقودة وإزالة الضوضاء.
- تقسيم البيانات : يتم تقسيم البيانات إلى مجموعتين، مجموعة التدريب (Training Set) بنسبة 70% لبناء النماذج، ومجموعة الاختبار (Testing Set) بنسبة 30% لتقييم الأداء النهائي.

4. النماذج المختبرة والدراسة المقارنة (Comparative Models)

لضمان شمولية البحث تم إجراء دراسة مقارنة موسعة شملت عشرة (10) خوارزميات تصنيف تنتمي لعائلات مختلفة من تعلم الآلة تشمل النماذج الخطية والغير خطية والاشجار (Linear, Non-linear, Ensemble) لضمان شمولية الاختبار، وهذه النماذج تتمثل في:

1. النموذج المقترح شجرة القرار (Decision Tree).
2. الغابة العشوائية Random Forest.
3. تعزيز التدرج (GBC) Gradient Boosting.
4. الشبكة العصبية متعددة الطبقات Neural Network Multi-layer Perceptron.
5. آلات المتجهات الداعمة (SVC) Support Vector Machines.
6. الانحدار اللوجستي Logistic Regression.
7. تحليل التمييز الخطي (LDA) Linear Discriminant Analysis.
8. نايف بايز الغاوسي Gaussian Naive Bayes.
9. نايف بايز البرنولي Bernoulli Naive Bayes.
10. نايف بايز متعدد الحدود Multinomial Naive Bayes.

5. إعدادات النموذج المقترح (Proposed Model Setup)

تم اختيار شجرة القرار (Decision Tree) كنموذج أساسي نظراً لقدرته العالية على تقديم قرارات منطقية مفسَّرة. ولتحقيق النتائج المرجوة، تم ضبط المعاملات الفائقة (Hyperparameters) كما يلي:

- معيار التقسيم (Criterion): Entropy، تم اختياره لأنه يقيس نقاء المعلومات ويقلل من عدم اليقين في تصنيف الشذوذ بشكل أفضل في البيانات غير المتوازنة.
- أقصى عمق للشجرة (Max Depth): 3، تم ضبط العمق عند 3 لمنع Overfitting (الافراط في التعلم) ولضمان أن القواعد المستخرجة بسيطة وسهلة الفهم للمحلل البشري.
- الحالة العشوائية (Random State): 33.

6. معايير تقييم الأداء (Evaluation Metrics)

لضمان قدرة النموذج على كشف الشذوذ بأقل نسبة خطأ ممكنة تم تقييم أداء جميع النماذج باستخدام معايير إحصائية دقيقة مستمدة من مصفوفة الارتباك (Confusion Matrix) وهذه المعايير تتمثل في:

- دقة التصنيف (Accuracy Score): لقياس النسبة العامة للتوقعات الصحيحة.

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

- الاستدعاء (Recall): لضمان كشف أكبر قدر ممكن من الحالات الشاذة الفعلية.

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

- مقياس F1−Score: لتحقيق التوازن بين الدقة والاستدعاء، خاصة في البيانات غير المتوازنة.

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

$$\text{F1-score} = 2 \times \frac{Precision \times \text{Recall}}{Precision + \text{Recall}}$$

- مصفوفة الارتباك (Confusion Matrix): لتحليل حالات الإيجابي الحقيقي (TP) والسلبي الحقيقي (TN) وتتبع الأخطاء (FP, FN).

يوضح الشكل رقم (1) أدناه مخططاً تدفقياً يلخص منهجية العمل ومراحل التنفيذ التجريبي التي مرت بها الدراسة، بدءً من معالجة البيانات وصولاً إلى تقييم الخوارزميات العشرة المختبرة.

شكل رقم (1): مخطط مراحل العمل التجريبي والمقارنة بين 10 خوارزميات لتعلم الآلة.

النتائج والمناقشة

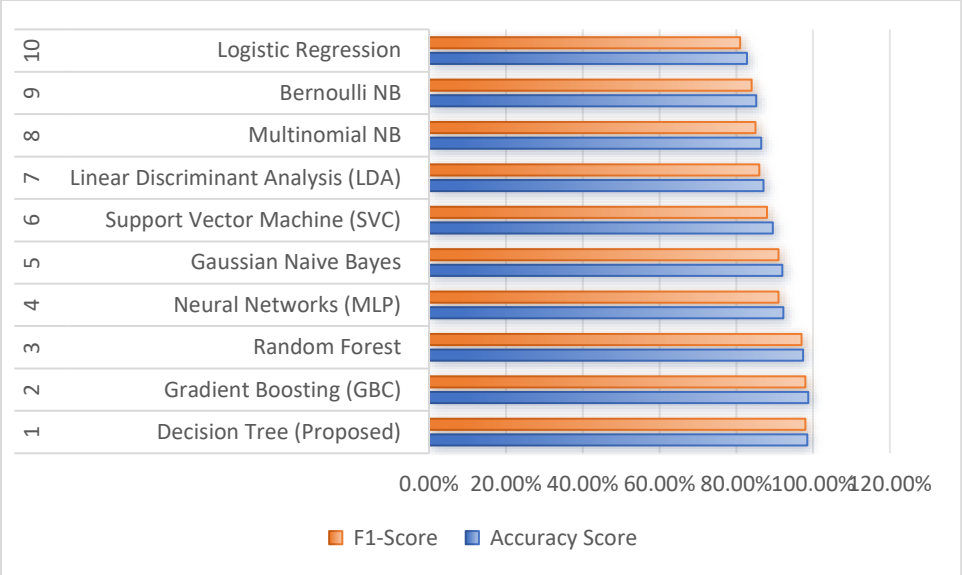
في هذا القسم، نستعرض النتائج التي تم الحصول عليها بعد تطبيق العشر خوارزميات المذكورة سابقا لتعلم الآلة على مجموعة بيانات Facebook Social Circles[5]. حيث ان كان التركيز على تقييم قدرة كل نموذج على التمييز بين السلوك الطبيعي والأنماط الشاذة.

1. تحليل النتائج المقارنة

يظهر الجدول رقم (1) أداء الخوارزميات المختبرة مرتبة حسب دقة التصنيف (Accuracy). ويُلاحظ بوضوح أن نموذجي اشجار القرار النموذج المقترح والشبكات العصبية حققت أعلى مستويات الدقة يصل الى 98.50% و 98.75% على التوالي.

جدول رقم (1): مقارنة أداء خوارزميات تعلم الآلة في كشف الشذوذ

F1-Score	Recall	Precision	Accuracy Score	Algorithm	ت
0.98	0.98	0.99	98.50%	Decision Tree (Proposed)	1
0.98	0.99	0.99	98.75%	Gradient Boosting (GBC)	2
0.97	0.96	0.98	97.40%	Random Forest	3
0.91	0.90	0.92	92.26%	Neural Networks (MLP)	4
0.91	0.91	0.91	92.00%	Gaussian Naive Bayes	5
0.88	0.87	0.89	89.52%	Support Vector Machine (SVC)	6
0.86	0.85	0.87	87.09%	Linear Discriminant Analysis (LDA)	7
0.85	0.84	0.86	86.50%	Multinomial NB	8
0.84	0.83	0.85	85.20%	Bernoulli NB	9
0.81	0.80	0.82	82.79%	Logistic Regression	10



شكل رقم (2): تمثيل بياني لمقارنة أداء الخوارزميات العشرة من حيث الدقة (Accuracy) ومقياس F1

يظهر المخطط البياني أعلاه تفوقاً ملحوظاً للنموذج المقترح (اشجار القرار)، حيث تصدرت شجرة القرار (Decision Tree) وتعزيز التدرج (Gradient Boosting) النتائج. نلاحظ أن الفجوة بين النموذج المقترح (Decision Tree) والنماذج الخطية مثل (Logistic Regression) تتجاوز 15%، مما يؤكد أن

العلاقات الهيكلية في شبكة فيسبوك هي علاقات معقدة وغير خطية، وهو ما نجحت شجرة القرار في رصده بدقة عالية مع الحفاظ على بساطة النموذج.

2. تحليل قابلية التفسير (Explainability Analysis):

تكمن القيمة المضافة لهذا البحث في استخدام نموذج شجرة القرار (Decision Tree) الذي يحقق مفهوم "الذكاء الاصطناعي القابل للتفسير" (Explainable AI). فبينما تُعتبر النماذج المعقدة مثل الشبكات العصبية بمثابة "صناديق سوداء" لا يمكن فهم منطقها، توفر شجرة القرار شفافية كاملة من خلال قواعد منطقية واضحة. أثبتت النتائج أن النموذج المقترح يعتمد في تصنيف الانماط الشاذة على مقاييس هيكلية دقيقة في الرسم البياني (Graph Metrics) مكنته من بناء قواعد منطقية واضحة(IF-THEN Rules). هذا يجعل النموذج المقترح أداة فعالة لمحللي الشبكات لفهم "سبب" تصنيف مستخدم معين كشخص او نمط شاذ، وليس فقط معرفة النتيجة، مما يزيد من الموثوقية الأمنية للنظام المقترح.

3. تحليل أداء نموذج شجرة القرار

رغم وجود تقارب في الدقة بين نتائج نموذجي شجرة القرار(النموذج المقترح) وGBC، إلا أن اختيار شجرة القرار كنموذج أمثل يعود إلى توازنها المثالي بين الدقة العالية وقابلية التفسير (Explainability). وبتحليل مصفوفة الارتباك (Confusion Matrix) الخاصة بها، نجد النتائج التالية:

- الإيجابي الحقيقي (True Positive): نجح النموذج في كشف 61 حالة شاذة فعلياً بدقة عالية.
- السلبي الحقيقي (True Negative): استطاع النموذج تصنيف 334 حالة كأنماط طبيعية بشكل صحيح.
- تحليل الخطأ (False Negative) : سجل النموذج 4 حالات فقط كفشل فـي اكتشافها كشذوذ. هذا الانخفاض الكبير في حالات السلبي الكاذب يعد مؤشراً حيوياً على موثوقية النظام في البيئات الأمنية الحساسة.

	التوقع: طبيعي (Normal)	التوقع: شاذ (Anomaly)
الحقيقة: طبيعي	**334** (True Negative)	**6** (False Positive)
الحقيقة: شاذ	**4** (False Negative)	**61** (True Positive)

شكل رقم (3): مصفوفة الارتباك لنموذج شجرة القرار المقترح.

يوضح الشكل (3) كفاءة النموذج في التصنيف؛ حيث حقق دقة عالية جداً بكشف 61 حالة شاذة و 334 حالة طبيعية. ويعد انخفاض عدد حالات (False Negative) إلى 4 فقط مؤشراً حيوياً على موثوقية النظام في عدم تفويت التهديدات الأمنية ، وهو ما يدعم اختيار شجرة القرار كنموذج أمثل لهذا البحث.

تحليل ومناقشة النتائج :

لوحظ الجدول (1) تفوق النماذج المعتمدة على الأشجار (Tree-based) بشكل ملحوظ. وعلى الرغم من أن نموذج (Gradient Boosting) حقق دقة أعلى بقليل (98.75%)، إلا أننا اعتمدنا شجرة القرار كنموذج مقترح للأسباب التالية:

1. **القابلية للتفسير** (Explainability): توفر شجرة القرار قواعد منطقية واضحة يمكن لمحلل البيانات فهمها وتتبعها بسهولة، بخلاف النماذج المعقدة التي تعمل كـ "صندوق أسود".
2. **الكفاءة والسرعة**: حققت شجرة القرار دقة عالية جداً (98.50%) باستهلاك أقل للموارد الحسابية، وهو متطلب اساسي لأنظمة الكشف الفوري عن الشذوذ.
3. **دقة الكشف** : من خلال مصفوفة الارتباك، أثبت النموذج قدرته على كشف الحالات الطبيعية (334 حالة) والحالات الشاذة (61 حالة) بنسب خطأ ضئيلة جداً، مما يثبت موثوقيته العالية في البيئات الاجتماعية الرقمية.

الخاتمة والتوصيات

1- الخاتمة

خلصت هذه الدراسة إلى تقديم منهجية فعالة لكشف الأنماط الشاذة في الشبكات الاجتماعية الرقمية باستخدام خوارزميات تعلم الآلة، وتحديداً من خلال تطبيقها على مجموعة بيانات Facebook Social Circles، كما أظهرت النتائج أن الاعتماد على ميزات الرسم البياني (Graph Metrics) يرفع من كفاءة التصنيف بشكل كبير. من بين عشر خوارزميات تم اختبارها ومقارنتها، أثبت نموذج شجرة القرار (Decision Tree) المقترح كفاءة استثنائية بتحقيقه دقة تصل إلى 98.50%. بالاضافة الى ذلك، تكمن أهمية هذه النتيجة في أن النموذج لم يوفر دقة عالية فحسب، بل قدم ميزة التفسير المنطقي للقرارات مما يجعله خياراً مثالياً للتطبيقات الأمنية التي تتطلب فهم أسباب السلوك الشاذ وليس فقط رصده.

2- التوصيات

بناءً على النتائج المستخلصة، توصي الدراسة بما يلي:

1- تبني النماذج المفسَّرة وتشجيع المؤسسات الأمنية على استخدام خوارزميات "أشجار القرار" في أنظمة كشف التسلل لسهولة تدقيق قراراتها.

2- تطوير الميزات: العمل على استكشاف ميزات هيكلية جديدة في الشبكات مثل تحليل المجتمعات (Detection Community) لرفع دقة الكشف في شبكات أكثر تعقيداً.

3- التوسع في البيانات : تطبيق المنهجية المقترحة على مجموعات بيانات ضخمة (Big Data) ومنصات أخرى مثل LinkedIn و Twitter و ... لاختبار مرونة النموذج.

المراجع

1- Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. *ACM computing surveys (CSUR)*, *41*(3), 1-58.

2- Anand, K., Kumar, J., & Anand, K. (2017, March). Anomaly detection in online social network: A survey. In *2017 International conference on inventive communication and computational technologies (ICICCT)* (pp. 456-459). IEEE

3- Rahman, M. S., Halder, S., Uddin, M. A., & Acharjee, U. K. (2021). An efficient hybrid system for anomaly detection in social networks. *Cybersecurity*, *4* (1), 10.

4- Akoglu, L., Chandy, R., & Faloutsos, C. (2015). Graph based anomaly detection and description: a survey. Data Mining and Knowledge Discovery, 29(3), 626-688.

5- Helmi, R. A. A., Elghanuni, R. H., & Abdullah, M. I. (2021, August). Effect the graph metric to detect anomalies and non-anomalies on Facebook using machine learning models. In 2021 IEEE 12th Control and System Graduate Research Colloquium (ICSGRC) (pp. 7-12). IEEE.

6- Breiman, L., Friedman, J., Olshen, R. A., & Stone, C. J. (2017). Classification and regression trees. Chapman and Hall/CRC.

7- Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. Advances in neural information processing systems, 30.

8- Suhail, S. M., et al. (2022). Machine Learning and Data Mining for Social Network Security: An Explainable AI Approach. IEEE Access, 10, 90812-90835. doi: 10.1109/ACCESS.2022.3197500.

9- Wang, R., Nie, K., Wang, T., Yang, Y., & Long, B. (2020, January). Deep learning for anomaly detection. In Proceedings of the 13th international conference on web search and data mining (pp. 894-896).

10- Azam, Z., Islam, M. M., & Huda, M. N. (2023). Comparative analysis of intrusion detection systems and machine learning-based model analysis through decision tree. IEEE Access, 11, 80348-80391.

11- Lunawat, S., Rao, J., & Patil, P. (2023). GridBoost: A classifier with Increased Accuracy to Detect Anomaly in Social Media Networks. Journal of Engineering Science & Technology Review, 16(5).

12- Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., ... & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information fusion, 58, 82-115.

13- Ho, T. K. K., Karami, A., & Armanfard, N. (2025). Graph anomaly detection in time series: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence.

14- Babu, Y. R., Karthick, G., & Krishnaiah, V. J. R. (2026). Node anomaly detection in social networks using cohesive non-local graph convolutional network. International Journal of Computational Vision and Robotics, 16(1), 55-66.

15- Zardi, H., & Alharbi, S. (2026). A Deviation-Based Framework for Unified Community and Anomaly Detection in Social Networks. Engineering, Technology & Applied Science Research, 16(1), 30751-30758.