

Syrian Arab Republic
Ministry of Higher Education
Syrian Virtual University

الجمهورية العربية السورية
وزارة التعليم العالي
الجامعة الافتراضية السورية

بحث مقدم لنيل شهادة ماجستير التأهيل والتخصص في إدارة التقنية PMTM

بعنوان

**دور تقنيات الذكاء الاصطناعي في إدارة أمن مواقع الويب: دراسة حالة تطبيقية للحماية
من الهجمات الخبيثة**

**The Role of Artificial Intelligence Technologies in Web Security
Management: A Case Study on Protection Against Malicious
Attacks**

إعداد:

الطالبة: واحه أيمن شعبان

Waha_214942

إشراف:

د. باسل يونس

F23

الإهداء

الحمد لله ع التمام والختام... الحمد لله الذي بلغني شرف هذا العلم وأعانني على إكماله.

أهدي نجاحي إلى من ربباني صغيرا وأولياني الرعاية والاهتمام...أمي وأبي العزيزين

أطال الله في عمرهما ومتعهما بالصحة والعافية...

زوجي الغالي د. أحمد إواهيم وابنتي فلذة كبدي ألين.. أجمل عطايا الله لي ...

إلى أستاذي الكريم المشرف على الوسالة الدكتور باسل يونس؛ لرعايته لي في نواستي

وبحثي بعلمه ووقته وتوجيهه المستمر وإرشاداته.

الشكر

الحمد لله الذي أنعم علينا بنعمة العلم ورفع مكانتنا بالعلم وعلما ما لم نكن نعلم، الحمد لله الذي أسبق علينا نعمه ظاهرة وباطنة..

قال رسول الله – صلى الله عليه وسلم -: " من لا يشكر الناس لا يشكر الله "، فإنه من دواعي سروري أن أخط هذه الكلمات لأتقدم بالشكر الجزيل والعرفان الجميل لكل من كان عوناً لي وسنداً في مسيرتي العلمية وأخص بالذكر:

التي جعل الله الجنة تحت أقدامها رمز الحب والعطف والحنان... أمي

إلى من علمني حب العلم والمثابرة والاجتهاد... أبي

إخوتي الغاليين... "م. فراس، د. دعاء، د. لوليا، أ. مها، د. أحمد"

زوجي الغالي كنت وستبقى خير سند وخير معلم.. لم تتركني أواجه التحديات بمفردي، أهديك هذا العمل تعبيراً عن شكري وامتناني لكل ما قدمته لي، وإيماناً مني بأن نجاحي هو نجاحنا معاً وأن وجودك بجانبني جعل كل هذا ممكناً..

ابنتي شمسي ونور أيامي "ألين".. لقد قاسمك هذا العمل طفولتك، وحرمتك من بعض أوجه اللعب والمرح.. أخط لك هذه الكلمات لكي تعلمي أن العلم هو من يرفع الانسان وتكن وصيتي لك بالعلم والصبر والاجتهاد والمثابرة.

أستاذي الفاضل المشرف على الرسالة الدكتور باسل يونس؛ بفضل نصائحك السديدة وملاحظاتكم البناءة، تمكنت من تجاوز العديد من التحديات والصعوبات التي واجهتني خلال فترة إعداد هذه الرسالة. لا يمكنني إلا أن أعبر عن شكري وتقديري لكل ما بذلتموه من جهد ووقت في سبيل تحقيق هذا العمل.

كما أتقدم بوافر الشكر والتقدير للجنة المناقشة الأفاضل الذين شرفوني بقبول مناقشة الرسالة. كما أتقدم بالشكر الجزيل والتقدير العميق إلى جميع أعضاء الهيئة التدريسية في الجامعة الافتراضية السورية.

إلى أرواح شهدائنا الأبرار... أهدي هذا العمل...

صفحة لجنة الحكم

الملخص

نتيجةً لتزايد وخطورة الهجمات السيبرانية التي تستهدف مواقع الويب والتي تُعرض بيانات المستخدمين للخطر وتلحق الضرر بسمعة المواقع مما يشكّل انتهاكاً لأمان تلك المواقع. تلعب تقنيات الذكاء الاصطناعي دوراً كبيراً في تعزيز إدارة أمن مواقع الويب، من خلال اكتشاف الهجمات الخبيثة ورسائل البريد الإلكتروني التصيدية.

يتمحور البحث حول تقييم أداء نماذج الذكاء الاصطناعي في اكتشاف وتصنيف الهجمات ورسائل البريد الإلكتروني على مواقع الويب باستخدام مجموعات بيانات مختلفة، تم تدريب النماذج للتعرف على الأنماط الخبيثة وتصنيفها ومن ثم تم تقييم أداءها في عملية الكشف، وفي المرحلة الأخيرة تم تطبيق هذه النماذج على موقع ويب افتراضي لاختبار أداؤها بشكل فعلي.

أظهرت نتائج البحث أن قيمة دقة نموذج التعلم العميق في كشف وتصنيف رسائل البريد الإلكتروني تتجاوز 97% مع نسبة ضياع أقل من 10%. إضافة إلى عدم وجود مشاكل الضبط الزائد (Overfitting). وبالنسبة للنتائج المتعلقة بكشف الهجمات السيبرانية، بلغت قيمة دقة نموذج التعلم الآلي بالمجمل حوالي 94% مع دقة 100% لتصنيف حركة المرور الطبيعية.

تم تطبيق نتائج الدراسة التحليلية على موقع ويب افتراضي، من خلال ربط نماذج الذكاء الاصطناعي مع موقع الويب، وأظهرت نتائج التطبيق فعالية النماذج في اكتشاف الهجمات ورسائل البريد الإلكتروني التصيدية التي تستهدف موقع الويب وتصنيف حركة المرور الطبيعية. تُمثل نتائج البحث إضافة قيمة لفهم كيفية استخدام الذكاء الاصطناعي في اكتشاف الهجمات السيبرانية وتعزيز أمن مواقع الويب.

كلمات مفتاحية: أمن مواقع الويب- الذكاء الاصطناعي- الهجمات السيبرانية.

Abstract

Due to the increasing and escalating severity of cyberattacks targeting websites, which jeopardize user data and damage website reputation, constituting a breach of website security, artificial intelligence technologies play a significant role in enhancing website security management by detecting malicious attacks and phishing emails.

The research focuses on evaluating the performance of artificial intelligence models in detecting and classifying attacks and emails on websites using different datasets. The models were trained to recognize and classify malicious patterns, and their performance in the detection process was then evaluated. In the final stage, these models were applied to a virtual website to test their performance in a real-world setting.

The results of research showed that the accuracy value of the deep learning model in detecting and classifying emails exceeds 97% with a loss rate of less than 10%. Additionally, there were no overfitting issues. As for the results related to cyberattack detection, the overall accuracy value of the machine learning model was around 94% with 100% accuracy for classifying normal traffic. The results of the analytical study were applied to a virtual website by linking artificial intelligence models to the website. The application results demonstrated the effectiveness of the models in detecting attacks and phishing emails targeting the website and classifying normal traffic. The research results represent a valuable addition to understanding how artificial intelligence can be used to detect cyberattacks and enhance website security.

Key words: Website security, Artificial intelligence, cybersecurity.

قائمة المحتويات

المحتويات

x.....	قائمة الجداول
xi.....	قائمة الأشكال
xiii.....	قائمة الملاحق
xiv.....	قائمة الاختصارات
1.....	الفصل الأول
1.....	الإطار العام للبحث
1.....	1-1 مقدمة
2.....	2-1 مشكلة البحث
3.....	3-1 هدف البحث
3.....	4-1 دراسات سابقة:
7.....	5-1 متغيرات البحث
7.....	6-1 فرضيات البحث
8.....	7-1 محددات البحث المكانية والزمنية:
8.....	8-1 مخطط الأطروحة:
10.....	الفصل الثاني
10.....	الهجمات السيبرانية على مواقع الويب
10.....	1-2 تمهيد:
10.....	2-2 أبرز الهجمات السيبرانية على مواقع الويب:
10.....	1-2-2 هجوم حجب الخدمة وحجب الخدمة الموزع:
11.....	2-2-2 هجمات كلمات المرور:
12.....	3-2-2 حقن لغة الاستعلام البنوية (SQL):
12.....	4-2-2 الاختراق عبر النصوص (Cross Site Scripting (XSS):
12.....	5-2-2 التنصت (Eavesdropping):
13.....	6-2-2 هجمات الوسيط (MitM Attacks):

13	7-2-2 التصيد (Phishing):
17	3-2 الذكاء الاصطناعي (Artificial Intelligent):
18	1-3-2 التعلم الآلي (Machine Learning):
19	2-3-2 التعلم العميق (Deep learning):
21	4-2 كشف الهجمات السيبرانية المعتمد على الذكاء الاصطناعي:
23	الفصل الثالث:
23	منهجية العمل
23	1-3 لمحة عامة:
23	2-3 أدوات البحث:
23	1-2-3 لغة البايثون (Python):
24	2-2-3 نوتبوك جوبيتر Jupyter Notebook:
24	3-3 المنهجية المقترحة لكشف رسائل البريد الإلكتروني التصيدي:
26	1-3-3 تحميل البيانات ذات الصلة:
27	2-3-3 مرحلة إظهار وتحليل البيانات:
29	3-3-3 المعالجة الأولية للبيانات:
30	4-3-3 مرحلة تحليل النصوص باستخدام معالجة اللغات الطبيعية NLP:
31	5-3-3 بناء نموذج الشبكة العصبية العميقة (Deep Neural Network):
32	4-3 المنهجية المقترحة لكشف الهجمات على مواقع الويب:
33	1-4-3 تحميل مجموعة البيانات الخاصة بالهجمات على مواقع الويب:
34	2-4-3 قراءة وتحليل مجموعة البيانات:
36	3-4-3 المعالجة الأولية للبيانات:
36	4-4-3 نموذج التعلم الآلي (شجرة القرار):
37	5-3 بارامترات تقييم الأداء:
37	1-5-3 معدل الخطأ (Error) والدقة (Accuracy):
38	2-5-3 الضبط والاستدعاء و F1:
39	3-5-3 مصفوفة الارتباك (Confusion Matrix):
40	6-3 الاستبيان:
40	1-6-3 مناقشة نتائج الاستبيان:

42.....	الفصل الرابع.....
42.....	النتائج والمناقشة.....
42.....	1-4 تمهيد:.....
42.....	2-4 تقييم أداء نماذج اكتشاف وتصنيف رسائل البريد الإلكتروني:.....
43.....	1-2-4 نتائج الدقة:.....
43.....	2-2-4 تابع الضياع:.....
44.....	3-2-4 مصفوفة الارتباك (Confusion Matrix):.....
45.....	4-2-4 الضبط والاستدعاء و F1:.....
46.....	3-4 تقييم أداء نماذج اكتشاف وتصنيف الهجمات على مواقع الويب:.....
48.....	4-4 الدراسة التطبيقية لكشف وحماية موقع ويب من الهجمات الخبيثة:.....
48.....	1-4-4 اختبار كشف وتصنيف رسائل البريد الإلكتروني:.....
51.....	2-4-4 اختبار كشف وتصنيف الهجمات:.....
54.....	الفصل الخامس.....
54.....	الاستنتاجات والتوصيات والمقترحات.....
54.....	1-5 الاستنتاجات:.....
55.....	2-5 التوصيات المستقبلية:.....
55.....	3-5 المقترحات:.....
56.....	المراجع:.....
	ملحق (أ) رابط
72.....	الاستبيان.....
	ملحق (ب) نتائج
73.....	الاستبيان.....

قائمة الجداول

الصفحة	العنوان	رقم الجدول
46	نتيجة بارامترات الأداء: الدقة والضبط والاستدعاء F1 و	(1-4)
47	نتائج تقييم الأداء لمختلف الخوارزميات	(2-4)

قائمة الأشكال

الصفحة	العنوان	رقم الشكل
11	هجوم حجب الخدمة وهجوم حجب الخدمة الموزع	(1-2)
14	رسائل البريد الإلكتروني التصيدي.	(2-2)
17	بعض مجالات الذكاء الاصطناعي	(3-2)
19	التعلم العميق	(4-2)
20	الخلية العصبية الصناعية	(5-2)
21	الشبكة العصبية المتكررة	(6-2)
25	منهجية العمل المقترحة لكشف رسائل البريد التصيدي	(1-3)
26	عينة من مجموعة البيانات الخاصة برسائل البريد الإلكتروني السليمة والتصيد	(2-3)
27	إظهار أول خمس أسطر من مجموعة البيانات	(3-3)
27	عدد العينات الكلي في مجموعة البيانات	(4-3)
28	الأصناف في مجموعة البيانات	(5-3)
28	مخطط للأصناف في مجموعة البيانات	(6-3)
29	نوع البيانات في مجموعة البيانات	(7-3)
29	مجموعة البيانات بعد حذف عمود غير ضروري	(8-3)
30	البيانات الفارغة في مجموعة البيانات	(9-3)
30	التأكد من إزالة البيانات الفارغة من مجموعة البيانات	(10-3)
30	قيم أصناف مجموعة البيانات بعد عمليات المعالجة الأولية	(11-3)

31	بنية الشبكة العصبية العميقة	(12-3)
32	المنهجية المقترحة لكشف الهجمات على مواقع الويب	(13-3)
34	جانب من خصائص مجموعة البيانات	(14-3)
35	توزيع العينات على أصناف الهجمات المستخدمة	(15-3)
39	مصفوفة الارتباك	(16-3)
43	دقة نموذج التعلم العميق بحالتي الاختبار والتدريب	(1-4)
44	تابع الضياع بحالتي الاختبار والتدريب	(2-4)
45	نتيجة مصفوفة الارتباك لتصنيف رسائل البريد	(3-4)
47	نتيجة مصفوفة الارتباك لكشف الهجمات	(4-4)
47	قيمة الدقة لكل صنف من أصناف مجموعة البيانات	(5-4)
49	اختبار إرسال رسالة بريد إلكتروني سليمة	(6-4)
49	نتيجة تصنيف رسائل البريد الإلكتروني السليمة من قبل نموذج الذكاء	(7-4)
50	اختبار إرسال رسالة بريد إلكتروني من نوع تصيد	(8-4)
50	نتيجة تصنيف رسالة البريد الإلكتروني من نوع تصيد من قبل نموذج الذكاء	(9-4)
51	اختبار توليد وكشف حركة المرور الطبيعية	(10-4)
52	اختبار توليد وكشف هجوم من Brute force	(11-4)
53	اختبار توليد وكشف هجوم من SQL Injection	(12-4)

قائمة الملاحق

الصفحة	العنوان	رقم الملحق
60	رابط الاستبيان	ملحق (أ)
60	نتائج الاستبيان	ملحق (ب)

قائمة الاختصارات

المعنى(عربي)	المعنى (انجليزي)	الاختصار
الذكاء الاصطناعي	Artificial Intelligent	AI
نظام كشف التسلل	Intrusion Detection System	IDS
الوحدات ذات البوابات المتكررة	Gated Recurrent Unit	GRU
خوارزمية الجار الأقرب	k-Nearest Neighbors	KNN
لغة الاستعلام البنوية	Structured Query Language	SQL
الشبكة الخاصة الافتراضية	Virtual Private Network	VPN
معالجة اللغات الطبيعية	Natural Language Processing	NLP
هجمات الوسيط	Man-in- the-Middle	MitM
الشبكة العصبية التكرارية	Recurrent neural network	RNN
الغابة العشوائية	Random Forest	RF

الفصل الأول

الإطار العام للبحث

1-1 مقدمة

تعتبر تقنيات الذكاء الاصطناعي (Artificial Intelligent) AI من أبرز التطورات التكنولوجية التي يشهدها العصر الحديث، حيث تمتلك القدرة على تحليل البيانات بشكل ذكي واتخاذ القرارات بناءً على الأنماط والتوقعات. تطبيقات الذكاء الاصطناعي تتخذ مسارًا متزايد الأهمية في مختلف المجالات، ومن بينها مجال أمن مواقع الويب، حيث تتعرض مواقع الويب بشكل متزايد للهجمات الخبيثة التي تستهدف سرقة المعلومات الحساسة، تعطيل الخدمات، أو تدمير البيانات.

إن أنظمة اكتشاف الاختراق التقليدية (IDS (Intrusion Detection System لا تستطيع اكتشاف جميع أنواع الهجمات والتصدي لها بشكل كامل، وذلك لأنها تقتصر على مراقبة الأنماط الثابتة في طلبات الويب [1]. فعندما يتم تشفير طلب الويب الضار بشكل بسيط، يمكنه تجاوز نظام اكتشاف الاختراق بسهولة وإحداث الضرر المطلوب ومن أجل مواجهة هذا التحدي المتزايد، يُعتبر دمج تقنيات الذكاء الاصطناعي في إدارة أمن مواقع الويب حلاً فعالاً ومبتكراً. حيث تتمتع خوارزميات التعلم الآلي بالقدرة على تعلم كمية كبيرة من الطلبات الآمنة والضارة ذات الأنماط المختلفة، ويمكنها التنبؤ بها بشكل فعال في البيئة التشغيلية [2].

يهدف هذا البحث إلى استكشاف وتطبيق دور تقنيات الذكاء الاصطناعي في تحسين أمن مواقع الويب، من خلال دراسة حالة تطبيقية تتناول كيفية استخدام هذه التقنيات في حماية المواقع من الهجمات الخبيثة، سوف يتم تناول ودراسة مجموعة بيانات تشمل أشهر الهجمات على مواقع الويب إضافة إلى مجموعة بيانات أخرى خاصة برسائل البريد الإلكتروني من نوع التصيد التي تحتوي على روابط مزيفة تستهدف سرقة أو تخريب البيانات الخاصة بالمستخدمين والإضرار بسمعة المؤسسة والأفراد كل على حدا. تسعى هذه الدراسة إلى تحليل الفوائد المحتملة لتطبيق تقنيات الذكاء الاصطناعي في مجال الأمن السيبراني، بالإضافة إلى تقديم نموذج عملي لتطبيق هذه التقنيات في مواجهة التهديدات الأمنية على

مواقع الويب. سيتم في هذا البحث استعراض الأدوات والتقنيات الرئيسية التي تُستخدم في تحقيق أمن مواقع الويب كلغة البايثون التي تعتبر إطار برمجي لتنفيذ وتقييم أداء خوارزميات الذكاء الاصطناعي، وتحليل كيفية تكامل تقنيات الذكاء الاصطناعي مع هذه الأدوات لتعزيز فاعلية الحماية. ستشمل الدراسة أيضاً تقييم النتائج المتوقعة والتحديات المحتملة في تطبيق هذه التقنيات، بالإضافة إلى اقتراحات لتحسين أداء أمن المواقع الإلكترونية باستخدام الذكاء الاصطناعي.

مراحل إنجاز هذا البحث، في المرحلة أولى سيتم إلقاء الضوء على دور تقنيات الذكاء الاصطناعي في مجال إدارة أمن مواقع الويب من جهة، وكيفية تحسين أمن تلك المواقع بمساعدة هذه التقنيات الواعدة والفعالة من جهة أخرى. أما المرحلة الثانية من هذا البحث ستشمل الاستفادة من الاستنتاجات التي يتم التوصل إليها في الدراسة التحليلية من أجل تطبيقها عملياً على موقع ويب افتراضي. حتى نصل في النهاية إلى رؤية متكاملة لإدارة أمن مواقع الويب نظرياً وعملياً، والاستفادة من هذه الرؤية كدراسة نموذجية من أجل تحقيق أمن مواقع الويب.

1-2 مشكلة البحث

تتمثل مشكلة البحث في الحاجة إلى فهم أفضل لدور تقنيات الذكاء الاصطناعي في إدارة أمن مواقع الويب وتطبيقاتها العملية لمكافحة الهجمات الخبيثة. على الرغم من التطورات الكبيرة في مجال الأمن السيبراني، إلا أن التهديدات الخبيثة لا تزال تتطور بسرعة وتتغير باستمرار. لذا، يتطلب ذلك بحثاً دقيقاً لفهم كيفية تكامل تقنيات الذكاء الاصطناعي مع استراتيجيات الأمن الحالية وإيجاد حلول فعالة لمواجهة هذه التحديات. ويمكن تلخيص مشكلة البحث الأساسية بالنقاط التالية:

- 1- ازدياد وتيرة الهجمات الخبيثة على مواقع الويب: تستهدف هذه الهجمات سرقة البيانات الحساسة، تعطيل الخدمات، أو تدمير البيانات.
- 2- قصور أنظمة اكتشاف الاختراق التقليدية: تعتمد هذه الأنظمة على مراقبة الأنماط الثابتة في طلبات الويب، مما يجعلها غير فعالة في اكتشاف الهجمات المشفرة أو الهجمات التي تُستخدم فيها تقنيات جديدة.
- 3- سهولة تجاوز أنظمة اكتشاف الاختراق التقليدية: يمكن للهجمات الخبيثة تجاوز هذه الأنظمة بسهولة من خلال تشفير طلبات الويب بشكل بسيط.

1-3 هدف البحث

هدف البحث هو تحليل وفهم دور تقنيات الذكاء الاصطناعي في إدارة أمن مواقع الويب، ودراسة تطبيقاتها العملية في مكافحة الهجمات الخبيثة. يهدف البحث أيضاً إلى تقديم حلول فعّالة لتحسين أمن مواقع الويب باستخدام تلك التقنيات. تتمثل أهمية هذا البحث في:

- تعزيز الفهم لكيفية استخدام التقنيات الذكاء الاصطناعي في تعزيز أمان مواقع الويب والحد من التهديدات السيبرانية.
- تقديم تصور عن الفوائد المحتملة والتحديات المرتبطة بتطبيق تقنيات الذكاء الاصطناعي في مجال أمن مواقع الويب.
- تقييم أداء نماذج تقنيات الذكاء الاصطناعي المستخدمة لمواجهة التحديات الأمنية الخاصة بمواقع الويب.
- توفير إطار عملي لتطبيق تقنيات الذكاء الاصطناعي في اكتشاف الهجمات بشكل فعّال على موقع ويب افتراضي.

من خلال تحقيق هذه الأهداف، يمكن للبحث أن يساهم في تعزيز الأمن السيبراني وحماية المعلومات على الويب، وبالتالي تحسين تجربة المستخدمين وبناء الثقة في استخدام مواقع الويب.

1-4 دراسات سابقة:

1- Web Server Attack Detection using Machine Learning (2020):

قدّم البحث نموذجاً يعتمد على التعلم الآلي للكشف عن التسلل باستخدام سجلات خادم الويب. يكتشف هذا النموذج ما إذا كان سجلاً معيناً طبيعياً أم سجل هجوماً، كما يحدد نوع الهجوم. يتم إنشاء سجلات خادم الويب وجمعها عن طريق إنشاء شبكة خاصة باستخدام خادم Apache WAMP [3].

2- Machine Learning based Intrusion Detection System for Web-Based Attacks (2020):

قام الباحثين بالتركيز على تحديد الأسباب الجذرية للنتائج الإيجابية الخاطئة والسلبية الخاطئة. وقد تم استخدام مجموعة بيانات CSIC 2010 HTTP. وأظهرت نتائج هذا البحث أن تطبيق استخراج

مجموعة الميزات المقترحة يؤدي إلى تحسين اكتشاف وتصنيف الهجمات القائمة على الويب لجميع خوارزميات التعلم الآلي التي تم اختبارها [4].

3- Machine learning-based intrusion detection system for detecting web attacks (2024):

حلل الباحثون بعض تقنيات التعلم الآلي التي تم اقتراحها في السنوات الأخيرة. كما تم إجراء تصنيفات متعددة للكشف عن السلوك غير الطبيعي في حركة مرور الشبكة. بُنيت النماذج وقيمت بناءً على مجموعة بيانات أنظمة الكشف عن التسلسل التي أصدرها المعهد الكندي للأمن السيبراني (CISE) في عام 2017، والتي تتضمن كلاً من الهجمات الحالية والهجمات التاريخية. أُجريت التجارب باستخدام خوارزميات شجرة القرار والغابة العشوائية والانحدار اللوجستي وطريقة naive bayes البسيطة (Gaussian) وخوارزمية AdaBoost والنهج المجمع لها. تم تقييم النماذج باستخدام مقاييس تقييم مختلفة مثل (accuracy) و (precision) و (recall) ومقياس F1 ومعدل الإيجابيات الخاطئة ومنحنى الـ ROC ومنحنى المعايرة [5].

4- Web Attack Intrusion Detection System Using Machine Learning Techniques (2024):

استخدم الباحثون مجموعة بيانات أنظمة الكشف عن التسلسل (CIC-IDS2017) التي يوفرها المعهد الكندي للأمن السيبراني لتقييم هجمات الويب. تم تقييم ثلاثة خوارزميات للتعلم الآلي في هذا البحث، وهي RF, KNN, NB وهي الهدف الأساسي من هذا البحث هو اقتراح خوارزمية تعلم آلي فعالة لنموذج أنظمة الكشف عن التسلسل لهجمات الويب. يقارن التقييم أداء الخوارزميات الثلاث بناءً على دقتها وفعالية اكتشافها لحركة مرور الشبكة التي تتعرض لهجوم ما. تشير النتائج إلى أن RF تفوقت على NB وعلى KNN من حيث متوسط الدقة المحقق خلال مرحلة التدريب. أما خلال مرحلة الاختبار، تفوقت خوارزمية KNN على RF وNB، حيث حققت معدل دقة متوسط يبلغ 99.4916%. ومع ذلك، حققت كل من RF و KNN معدل دقة متوسط يبلغ 100% مقارنة بالخوارزميات الأخرى. وأخيراً، تم تحديد خوارزميتي RF و KNN كأكثر الخوارزميات فعالية في اكتشاف هجمات الويب التي تستهدف أنظمة الكشف عن التسلسل [6].

5- Machine learning techniques applied to detect cyber-attacks on web applications(2015):

تناولت هذه الدراسة مواجهة هجمات الويب المبتكرة على مستوى التطبيق، حيث صنفت هذه الهجمات ضمن التهديدات الرئيسية والتحديات الأساسية لأمن الشبكات والإنترنت. تكمن المساهمة الرئيسية للمقال في اقتراح نهج التعلم الآلي لنمذجة السلوك الطبيعي للتطبيق واكتشاف الهجمات السيبرانية. يتكون النموذج من أنماط يتم الحصول عليها باستخدام تقنية التقسيم المعتمدة على الرسم البياني والبرمجة الديناميكية. يعتمد النموذج على المعلومات التي تم الحصول عليها من طلبات HTTP التي يرسلها العميل إلى خادم الويب. تم اعتماد مجموعة بيانات CSIC 2010 HTTP وحققت نتائج مرضية [7].

6- A Survey of Tools and Techniques for Web Attack Detection (2022):

يقسم البحث إلى قسمين رئيسيين: القسم الأول: يستعرض بعض الأدوات والتقنيات النموذجية لمراقبة واكتشاف هجمات الويب، التي تم تطويرها ونشرها عملياً. القسم الثاني: يستعرض التجارب وتقييم كفاءة النماذج القائمة على التعلم الآلي للكشف عن الهجمات الإلكترونية. وتؤكد النتائج التجريبية أن نموذج كشف هجمات الويب المعتمد على التعلم الآلي يحقق دقة كشف تصل إلى 99.57%، كما يتمتع النموذج بإمكانية التطبيق العملي [8].

7- Explainable machine learning for phishing feature detection (2023):

ساهم هذا البحث في مجال اكتشاف مواقع التصيد الاحتيالي من خلال اقتراح نموذج تعلم آلي قابل للتفسير. لا يوفر النموذج تنبؤات دقيقة حول مواقع التصيد الاحتيالي فحسب، بل قدم أيضاً تفسيرات للخصائص الأكثر ارتباطاً بمواقع التصيد الاحتيالي. ولتحقيق هذا الهدف، تم اقتراح نموذجاً جديداً لاختيار الميزات يعتمد على Lorenz zonoid، وهو امتداد متعدد الأبعاد لمعامل جيني. تم عرض المقترحات على مجموعة بيانات حقيقية تحتوي على ميزات التصيد الاحتيالي والمواقع الشرعية [9].

8- Assessment of Existing Cyber-Attack Detection Models for Web-Based Systems (2023):

يفحص هذا البحث النماذج الحالية للكشف عن هجمات الشبكة ويوصي بنماذج وتقنيات مناسبة للكشف عن هجمات الشبكة للأنظمة القائمة على الشبكة. من الواضح أن تكنولوجيا التعلم العميق توفر أداءً أفضل وقوة أكبر من التعلم الآلي التقليدي والتقنيات الأخرى غير المعتمدة على الذكاء الاصطناعي [10].

9- Deep Learning Technique-Enabled Web Application Firewall for the Detection of Web Attacks (2023):

هدفت هذه الدراسة إلى كيفية تعلم تقنية التعلم العميق الميزات وتستخرجها تلقائيًا دون تدخل بشري ويمكنها أيضًا التعامل مع البيانات الكبيرة ومتعددة الأبعاد بشكل أفضل من التقنيات الأخرى. توفر هذه الدراسة تحليلًا مقارنًا بين حركة مرور HTTP العادية وحركة مرور الهجوم، والذي يحدد المعايير والميزات التي تشير إلى وجود هجوم. تم تحليل ميزات مختلفة لمجموعات البيانات القياسية ISCX و CISC و CICDDoS، وتمت مقارنة حركة مرور الهجوم والمرور العادي من خلال أخذ معايير مختلفة بعين الاعتبار. كما تم تطوير نموذج معماري متعدد الطبقات لاكتشاف هجمات XSS و DDoS وحقن SQL باستخدام مجموعة بيانات تم جمعها من بيئة محاكاة. في هذا النموذج ذي الطبقات المستندة على الذاكرة طويلة وقصيرة الأجل (LSTM)، تكون الطبقة الأولى مسؤولة عن اكتشاف هجمات DDoS ودقتها 97.57٪، والطبقة الثانية تكتشف هجمات XSS وحقن SQL بدقة 89.34٪. يتم في البحث أولاً التحقق من حركة مرور HTTP ذات المعدل الأعلى ويتم تصفيتها، ثم تمريرها إلى الطبقة الثانية. يضيف جدار حماية تطبيقات الويب (WAF) طبقة إضافية من الأمان إلى تطبيق الويب عن طريق توفير فلتر على مستوى التطبيق لا يمكن تحقيقها بواسطة نظام جدار الحماية الشبكي التقليدي [11].

10- Deep Auto-Encoder Neural Network for Phishing Website Classification (2018):

قدّم البحث تقنية التشفير التلقائي العميق لحل مشكلة تصنيف مواقع التصيد الإلكتروني. تم الحصول على مجموعة البيانات من موقع UCI الذي يضم أشهر مجموعات بيانات التعلم الآلي. تتكون مجموعة

البيانات المستخرجة من 30 خاصية، والخاصية الحادية والثلاثون (المستهدفة) تمثل وجود تصيد احتيالي من عدمه. يعمل التشفير التلقائي الأول على استخراج الخصائص المهمة وتقليل أبعاد الميزات. ويقوم التشفير التلقائي الثاني أيضاً باستخراج الخصائص من مخرج التشفير التلقائي الأول ويقلل من أبعاد الميزات كذلك يتم تصنيف الخصائص المستخرجة باستخدام مصنف SoftMax ، يتم تكديس كل هذه الأجزاء وتدريبها باستخدام التعلم المُشرف. تظهر النتائج التجريبية أن الطريقة المقترحة تقدم أفضل النتائج مقارنة بالأعمال السابقة [12].

1-5 متغيرات البحث

متغيرات البحث هي العوامل التي يمكن أن تؤثر على نتائج الدراسة وتحليلاتها. المتغيرات المحتملة يمكن أن تشمل:

- نوع تقنيات الذكاء الاصطناعي: يمكن أن تختلف تقنيات الذكاء الاصطناعي المستخدمة في البحث، مثل تعلم الآلة، والتعلم العميق، حيث يؤثر اختيار هذه التقنيات على نتائج البحث.
- نوع التهديدات السيبرانية: يمكن أن تختلف الهجمات الخبيثة التي تستهدف مواقع الويب بشكل كبير، مما يمكن أن يؤدي إلى اختلاف في استراتيجيات الدفاع والتحليل التي يستخدمها البحث.
- حجم البيانات: يمكن أن يؤثر حجم ونوع البيانات المتاحة على قدرة البحث في تحليل التهديدات السيبرانية وتقديم الحلول المناسبة.

يساعد فهم هذه المتغيرات في توجيه البحث وتحديد العوامل الرئيسية التي يجب أخذها في الاعتبار للوصول إلى نتائج دقيقة وموثوقة.

1-6 فرضيات البحث

توجد عدة فرضيات ممكنة يمكن أن نقتربها في هذا البحث، ومن بين هذه الفرضيات:

- تطبيقات الذكاء الاصطناعي ستساهم في تحسين طرق اكتشاف وتصنيف الهجمات الخبيثة على مواقع الويب بشكل فعال.
- فرضية أن استخدام تقنيات الذكاء الاصطناعي في إدارة أمن مواقع الويب سيؤدي إلى تقليل معدلات الفشل في اكتشاف الهجمات مما يؤدي إلى تحسين أمان تلك المواقع.

- فرضية أن تكامل تقنيات الذكاء الاصطناعي مع أدوات الأمان التقليدية سيزيد من فعالية الحماية ويقلل من تأثير التهديدات السيبرانية.
- فرضية أن تقديم نماذج عملية لتطبيق تقنيات الذكاء الاصطناعي في حماية مواقع الويب سيساعد في تعزيز فهم كيفية تنفيذ هذه التقنيات في الواقع العملي.

1-7 محددات البحث المكانية والزمنية:

محددات البحث المكانية والزمنية قد تتضمن:

المكان: موقع ويب افتراضي كحالة نموذجية تقوم على تطبيق تقنيات الذكاء الاصطناعي التي تم تقييم أداءها وفعاليتها في الدراسة التحليلية، واستخدام هذه الأدوات عملياً لتقديم حماية متكاملة لأمن موقع الويب.

الزمان: يمكن أن يؤثر الزمان على نتائج البحث، حيث يمكن أن تتغير التهديدات السيبرانية وتطورات تقنيات الذكاء الاصطناعي مع مرور الوقت، وهذا يعني أن الدراسة قد تحتاج إلى التحديث والمراجعة المستمرة.

الجهات التي يمكنها الاستفادة من البحث: من أبرز الجهات التي يمكنها الاستفادة من هذا البحث مواقع الويب التابعة للمؤسسات العامة والخاصة والمنظمات والبنوك من أجل حماية البيانات الخاصة بها والحفاظ على سريتها.

1-8 مخطط الأطروحة:

تُنظَّم أجزاء الأطروحة كما يلي:

في الفصل الأول: الإطار العام للدراسة، مع توضيح مشكلة البحث وأهداف البحث وأهميته، واستعراض أبرز الدراسات المرجعية المتعلقة بموضوع البحث.

في الفصل الثاني: سيتم استعراض الدراسة النظرية حول أبرز الهجمات التي تستهدف مواقع الويب والطرق التقليدية للحد من هذه الهجمات سواء كانت هجمات خبيثة أو رسائل بريد تصيد. إضافة لذلك، يتم توضيح مفاهيم الذكاء الاصطناعي والتعلم الآلي والتعلم العميق ودورها في كشف وتصنيف الهجمات على مواقع الويب.

في الفصل الثالث من هذه الأطروحة: يتم توضيح منهجية البحث المقترحة، حيث يتم تقسيمها إلى منهجيتين، أولى خاصة برسائل البريد الإلكتروني والثانية خاصة بأشهر الهجمات على مواقع الويب، إضافة لذلك تم إجراء استبيان حول الهجمات السيبرانية وطرق التصدي لها ودور الذكاء الاصطناعي في كشف هذه الهجمات.

في الفصل الرابع: يتم استعراض نتائج البحث ومناقشتها وتقييم بارامترات الأداء لنماذج الذكاء الاصطناعي المستخدمة لكشف الهجمات السيبرانية. إضافة لذلك يتم تطبيق وتنفيذ نتائج الدراسة التحليلية على موقع ويب افتراضي واختبار نماذج الذكاء الاصطناعي المقترحة بشكل حقيقي.

في الفصل الخامس: يتم تقديم الاستنتاجات النهائية التي هي خلاصة هذا البحث ونتيجته وتبيان أهميته في تحقيق الأمن والإدارة لمواقع الويب باستخدام الذكاء الاصطناعي، وأخيراً، يتم تقديم التوصيات المستقبلية والمقترحات لتطوير هذا العمل وتحسينه.

الفصل الثاني

الهجمات السيبرانية على مواقع الويب

1-2 تمهيد:

تُعرّف الهجمات السيبرانية بأنها شكل من أشكال الهجمات التي تستخدمها الجماعات أو الأفراد، والتي تستهدف أنظمة الكمبيوتر والبنية التحتية والشبكات أو الأجهزة الشخصية لتحقيق نتائج ضارة [13]، يتميز الهجوم السيبراني بأنه مجهول الهوية غالباً ويستهدف الشبكات والأنظمة الضعيفة إما لتعديل أو سرقة أو تدمير أهدافه. أصبحت العديد من المنظمات عرضة للخطر نتيجة لطبيعة الإنترنت المتاحة والسهولة الوصول إليها، وبالتالي فهي تشكل عامل جذب كبير لتهديدات الأمن السيبراني [14]. لذلك، أصبح من الضروري تأمين وحماية المعلومات والأنظمة من خلال تقنيات مختلفة مثل جدران الحماية والمصادقة والتشفير وأنظمة كشف هجمات الكمبيوتر وغيرها من حلول البرامج والأجهزة. هناك نوعان رئيسيان من الهجمات بشكل عام: موجهة وغير موجهة. في الهجمات الموجهة، يركّز المهاجمون على مؤسسات معينة بهدف إحداث ضرر أكبر مقارنة بالهجوم غير الموجه. أما الهجمات غير الموجهة، فيستهدف المهاجمون فيها مجموعة واسعة من المستخدمين والأجهزة، مستفيدين من انفتاح الإنترنت.

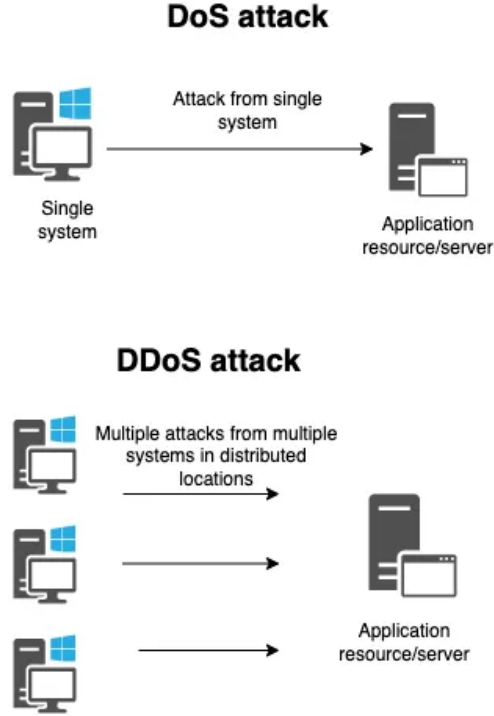
2-2 أبرز الهجمات السيبرانية على مواقع الويب:

لفهم كيفية التصدي بنجاح للهجمات السيبرانية، نحتاج إلى فهم كل من أشكال الهجمات المختلفة والمراحل التي يتم تنفيذ الهجوم من خلالها. لذلك سنتطرق إلى أشهر هذه الهجمات وهي:

2-2-1 هجوم حجب الخدمة وحجب الخدمة الموزع:

تهدف هجمات منع الخدمة (DoS) بشكل أساسي إلى إغراق موارد الأنظمة والشبكات بحيث لا يتم الاستجابة لطلبات الخدمة [15]. في هجوم منع الخدمة الموزع (DDoS)، تتعرض أجهزة الاستضافة لبرامج يسيطر عليها المهاجمون. وبالتالي، تصبح الموارد غير متوفرة للمستخدمين حيث تتعطل خدمة المضيف المرتبطة بالإنترنت كما يوضح الشكل (1-2).

منع هجمات حجب الخدمة (DoS) صعب للغاية بسبب التحدي المتمثل في التفريق بين التهديدات الصالحة وغير الصالحة من الطلبات في حركة مرور النظام أو الشبكة، وذلك لأن نفس البروتوكول والمنفذ يستخدمان في تقديم الطلبات.



الشكل (1-2): هجوم حجب الخدمة وهجوم حجب الخدمة الموزع

2-2-2 هجمات كلمات المرور:

كلمات المرور هي النهج الشائع لمصادقة المستخدم لمنع الهجمات. ومع ذلك، يمكن للمهاجمين فك تشفير كلمات المرور أو الحصول عليها ببساطة عن طريق الوصول إلى قواعد البيانات. برامج التجسس (sniffers) هي من الطرق التي يمكن أن ينفذها المهاجمون [16]. إضافة إلى القوة القاهرة (Brute Force) وهي طريقة ينفذها المهاجمون للحصول على كلمات المرور، تتضمن هذه الطريقة قيام المهاجمين بتجربة كلمات مرور مختلفة على أمل أن ينجحوا في الدخول إلى النظام أو الشبكة. يمكن تحقيق الوقاية من خلال إجراء تغييرات منتظمة لكلمة المرور وإضافة تعقيد إلى أنماط كلمة المرور.

2-2-3 حقن لغة الاستعلام البنيوية (SQL):

هجوم SQL هو أسلوب يستخدم فيه المهاجم رمزًا ضارًا للوصول إلى المعلومات عن طريق معالجة قاعدة البيانات الخلفية. المعلومات التي يستهدفها المهاجمون قد تكون تفاصيل حول المؤسسة المستهدفة أو تفاصيل عن عملاء الشركة أو مستخدمي أنظمة الشركة وشبكاتهما مما يؤدي إلى عرض غير مشروع للمعلومات الحساسة أو حذفها أو تعديلها [17].

يمكن منع حقن SQL عن طريق تنفيذ التحقق من صحة المدخلات والذي من شأنه تحديد المدخلات غير القانونية، ويتم اعتماد جدران الحماية لإزالة حقن SQL. ويمكن أيضًا استخدام عملية التعرف على عناوين IP المشبوهة وعلى التوقعات الرقمية في تحديد وحظر حقن SQL [18].

2-2-4 الاختراق عبر النصوص (Cross Site Scripting (XSS):

هو اختراق شائع حيث يقوم هذا النوع بحقن تعليمة برمجية خبيثة في مواقع صحيحة أو تطبيقات ويب. يحدث هذا النوع من الهجمات عندما يقوم المهاجمون بإدخال JavaScript أو كود إلى قواعد بيانات مواقع الويب، مما يجعل المستخدمين يقومون بتنزيلها. عند تنفيذ الضحية للسكريبت، يتمكن المهاجم من الوصول إلى ملفات تعريف ارتباطه (cookies) [19].

2-2-5 التنصت (Eavesdropping):

يُعرف هذا النوع من الهجوم أيضًا بأسماء أخرى مثل التجسس. يعتمد المهاجمون بشكل أساسي على اختراق بيانات الأجهزة الرقمية باستخدام شبكات غير آمنة لاستقبال البيانات وإرسالها. يتميز هذا الهجوم بحقيقة أنه لا يظهر أي نشوهات في عمليات الإرسال وبالتالي يصعب اكتشافه. يهدف المهاجمون الذين يستخدمون هذه الطريقة إلى الوصول إلى معلومات خاصة وحساسة مثل أرقام بطاقات الائتمان وكلمات المرور التي يتم تمريرها عبر الشبكات. يمكن إدخال برامج التجسس (Sniffers) لتنفيذ هذه الهجمات والسيطرة على البيانات المرسل [20]. يتم تحقيق الوقاية من هذه الهجمات من خلال برامج مكافحة الفيروسات وتنفيذ جدران الحماية واستخدام شبكات افتراضية خاصة (VPNs) والتشفير وعدم نقل البيانات الخاصة على الشبكات العامة.

2-2-6 هجمات الوسيط (MitM Attacks):

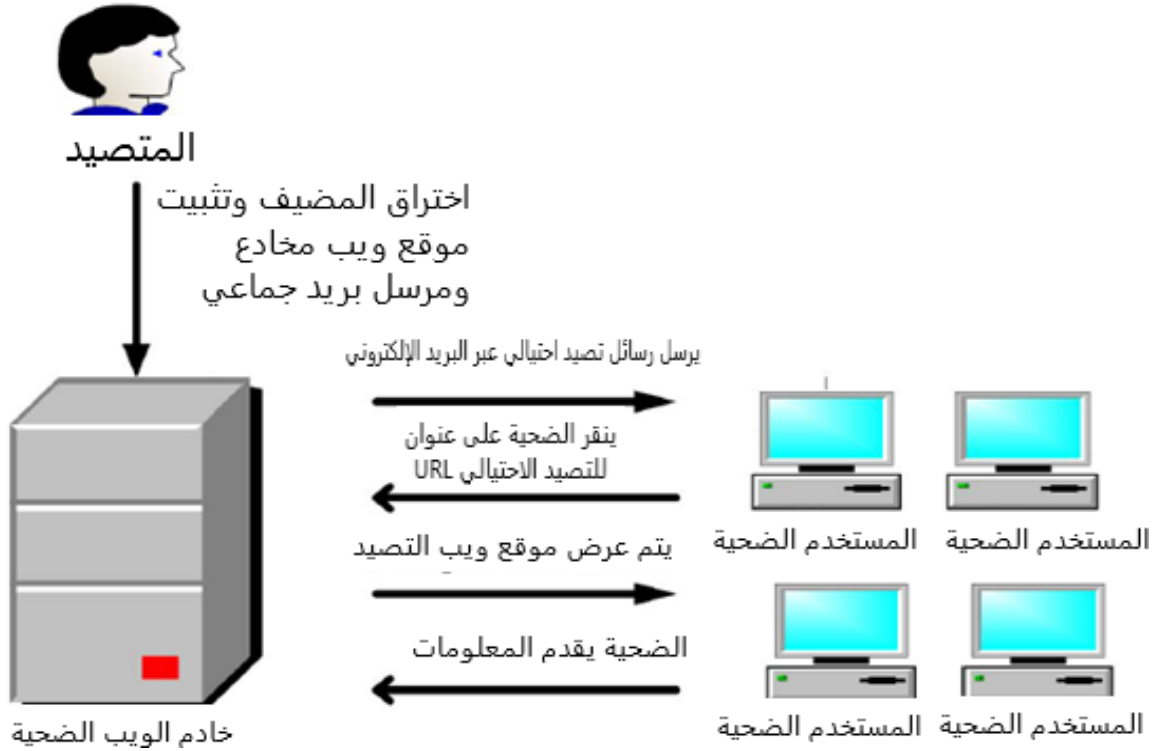
تحدث هجمات الوسيط (MitM) عندما يعطل طرف ثالث الاتصالات بين الخوادم والعملاء. يتمكن المهاجم من الحصول على المعلومات من خلال انتحال صفة كل من الخادم والعميل، وبالتالي يكون الخطر هو قدرة المهاجم على اعتراض المعلومات الموجهة لشخص آخر على الشبكة واستقبالها وإرسالها. في هذه الهجمات، يتم التدخل في عمليات تبادل المعلومات والعمليات المعاملاتية والاتصالات التي تتم في وقت معين. هناك أنواع مختلفة من الهجمات في هذه الحالة مثل انتحال IP واختطاف جلسات على الشبكة [21]. ولمنع هذا المستوى من الهجوم، يمكن تنفيذ أنظمة لكشف التسلل حيث أنها تعطي تنبيهات عند تنفيذ عمليات التسلل. بالإضافة إلى ذلك، تعد شبكات VPN مفيدة أيضاً في منع مثل هذه الهجمات لأنها تساعد في إنشاء المزيد من طبقات الأمان عند الوصول إلى الطبقات السرية لشبكة الشركة، خاصة عند القيام بذلك عبر اتصالات لاسلكية.

2-2-7 التصيد (Phishing):

يشتمل هذا النوع من الهجوم على قيام المهاجمين بإرسال رسائل بريد إلكتروني بشكل احتيالي تبدو مصادرهما وكأنها من مصادر صحيحة. يهدف هذا الهجوم إلى الحصول على بيانات الاعتماد والبيانات الشخصية. يعتبر التصيد هجوماً للهندسة الاجتماعية حيث تحتوي رسائل البريد الإلكتروني على روابط تشعبية مضمنة تطلق البرامج الضارة في شبكة النظام، أو تقود المستخدمين في بعض الأحيان إلى مواقع الويب التي تطلق البرامج الضارة في الشبكة [22].

تُعتبر رسائل البريد الإلكتروني المُخدعة إحدى أساليب الاحتيال وتُستخدم لسرقة البيانات الشخصية والمهمة للمستخدمين، حيث يتلقى المستخدمون رسالة بريد إلكتروني وهمية من عناوين مخادعة والتي تبدو أنها تنتمي إلى أعمال تجارية مشروعة وحقيقية في محاولة لسرقة المعلومات الشخصية للمتلقى. هذا العمل يشكل خطراً على خصوصية العديد من المستخدمين، وبالتالي يعمل الباحثون بشكل مستمر على إيجاد أدوات للكشف عن هذا النوع من الرسائل الإلكترونية وتطوير القائم منها. التصنيف هو أحد الطرق المتبعة في تقنيات التنقيب عن البيانات التي يمكن استخدامها بشكل فعال للكشف عن رسائل البريد الإلكتروني المخدعة.

تصنّف رسائل البريد الإلكتروني الاحتيالية (Phishing) كرسائل غير مرغوب فيها (Spam). حيث يتلقى المستخدمون رسائل بريد إلكتروني تدّعي أنها من شركة أو بنك شرعي، وتطلب من المستخدم اتّباع رابط مضمّن. سيقوم الرابط بإعادة توجيه المستخدم إلى موقع ويب مزيف يطلب معلومات سرية، مثل أسماء المستخدمين أو كلمات المرور أو أرقام بطاقات الائتمان... إلخ [23].



الشكل (2.2): رسائل البريد الإلكتروني التصيدي

الشكل (2.2) يوضّح دورة حياة تقنية التصيد الاحتيالي. حيث تبدأ العملية بإرسال رسائل بريد إلكتروني إلى صناديق الوارد الخاصة بالأفراد المستهدفين في محاولة لجعلهم يتابعون رابطاً مضمناً. بهذا المعنى، يشبه التصيد الاحتيالي عبر الإنترنت الصيد التقليدي؛ حيث يستخدم الصياد في أسلوب الصيد التقليدي طُعماً وصدارة لصيد الأسماك، بينما يرسل مرسل التصيد الاحتيالي أكبر عدد ممكن من رسائل البريد الإلكتروني في محاولة لإقناع أكبر عدد من المستلمين "بالتقاط" الطعم واتّباع الرابط المضمّن. عادةً يبدو البريد الإلكتروني شرعياً وممكن أن يحتوي على شعار شركة لمؤسسة مالية شهيرة وعنوان إرسال للشركة الشرعية. سيظهر الرابط في البريد الإلكتروني والمظهر أيضاً شرعياً للوهلة

الأولى. يريد المُصطاد أن يبدو الطعم أصليًا للغاية لدرجة أن الضحية ستقوم باتباع الرابط المزيف من دون تفكير [24].

1-7-2-2 التقنيات التقليدية للكشف عن رسائل البريد الإلكتروني الاحتيالي (Phishing):

لقد حظي الكشف عن رسائل البريد الإلكتروني الاحتيالية (Phishing) باهتمام كبير مؤخرًا نظرًا لتأثيرها على أمن المستخدمين. لذلك، تم استخدام العديد من التقنيات للكشف عن رسائل البريد الإلكتروني الاحتيالية منها:

1- فترة القائمة السوداء:

توفّر تقنية فلتر القائمة السوداء (blacklist filtering) الحماية على مستوى الشبكة من خلال تصنيف رسائل البريد الإلكتروني المستلمة بناءً على عنوان المرسل أو عنوان IP أو عنوان DNS. يتم استخراج هذه التفاصيل من ترويسة البريد الإلكتروني ومقارنتها بقائمة محددة مسبقًا، فإذا تطابقت أي من هذه البيانات مع القائمة سيتم رفض البريد الإلكتروني. وبالتالي، تقوم هذه التقنية بتصنيف رسائل البريد الإلكتروني الاحتيالية (Phishing) لتوفير الأمان على مستوى الشبكة. يُعد مقدمو خدمات الإنترنت (ISP) هم المنظمة المسؤولة عن توفير وتنفيذ هذه الفلتر.

2- فترة القائمة البيضاء:

توفّر فلتر القائمة البيضاء الحماية على مستوى الشبكة أيضًا، ولكن على عكس القوائم السوداء، تقارن هذه التقنية بيانات البريد الإلكتروني بقائمة محددة مسبقًا تحتوي على عناوين IP ثابتة لنطاقات شرعية وعناوين IP مشروعة. في هذا الصدد، يُسمح فقط بالبريد الإلكتروني الذي تتطابق بياناته مع القائمة بالوصول عبر الشبكة إلى الصندوق الوارد الخاص بالمستخدم [25].

يتم تضمين عناوين البريد الإلكتروني وعناوين IP في القائمة البيضاء إذا كانت تنتمي إلى مستخدمين أو شركات شرعية وافقت على إضافة عناوينها إلى هذه القائمة. سيتم تصنيف رسائل البريد الإلكتروني التي تتطابق بياناتها مع هذه القائمة فقط على أنها شرعية بناءً على هذا الفلتر، بينما تعتبر رسائل البريد

الإلكتروني الأخرى تصيداً احتياليًا ويمنع وصولها إلى الشبكة، ولذلك يُطلق على هذا الفلتر أيضًا "مصنف رسائل البريد الإلكتروني الشرعية".

3- فترة مطابقة الأنماط:

تستخدم تقنية مطابقة الأنماط (Pattern Matching) فلتر رسائل البريد الإلكتروني بناءً على أنماط محددة، بما في ذلك الكلمات والسلاسل النصية ومجموعات الأحرف المذكورة في محتوى البريد الإلكتروني أو عنوانه أو المرسل. يبحث الفلتر عبر البريد الإلكتروني عن هذه الأنماط المحددة لتصنيف البريد الإلكتروني إلى تصيد احتيالي أو شرعي.

على الرغم من أن هذه التقنية توفر الحماية على مستوى الشبكة وتعطي نتائج قيمة، إلا أنها لا تزال تحتوي على نتائج خاطئة وذلك بسبب العدد الهائل من رسائل البريد الإلكتروني الواردة والتي قد تتضمن كلمات أو سلاسل نصية محظورة ولكن لا ينبغي منعها.

4- التحقق من صحة البريد الإلكتروني:

التحقق من صحة البريد الإلكتروني (Email Verification) هي طريقة مصادقة على مستوى المستخدم تتطلب التحقق من كل من المرسل والمستقبل. بمجرد قبول المرسل لرسالة الإشعار، يتم اعتماد البريد الإلكتروني وتصنيفه على أنه شرعي للسماح له بالوصول إلى صندوق واردة المستقبل. خلاف ذلك، يُعتبر البريد الإلكتروني تصيداً احتياليًا وبالتالي يُمنع من الوصول إلى الصندوق الوارد [26].

5- فترة كلمة المرور:

توفر فلتر كلمة المرور الحماية أيضًا من خلال المصادقة على مستوى المستخدم، يسمح استخدام هذه الفلتر باستقبال أي بريد إلكتروني في سطر الموضوع أو عنوان البريد الإلكتروني أو حقل الرأس أو في أي جزء من البريد الإلكتروني فقط إذا تمكنت الفلتر من اكتشاف كلمة المرور المحددة. لذلك، إذا لم تتمكن الفلتر من العثور على كلمة المرور أو اكتشفت كلمة مرور خاطئة، فسيتم رفض البريد الإلكتروني. لا يتم إنشاء هذه الكلمات المرور افتراضياً، وبالتالي يتعين على المستخدمين الجدد لهذه الفلتر الدخول في محادثة مع بعضهم البعض لتعيين كلمة مرور وتفعيلها ثم يتم تصنيفها على أنها شرعية

بواسطة الفلتر. لا يزال هذا النوع من الفلاتر يعاني من قصور في فقدان بعض رسائل البريد الإلكتروني المشروعة إذا لم يتم التعرف على كلمة المرور، بالإضافة إلى أن العملية تستغرق وقتاً [27].

2-3 الذكاء الاصطناعي (Artificial Intelligent):

الذكاء الاصطناعي هو فرع من فروع علوم الحاسب يهدف إلى تعزيز قدرة الآلات والحواسيب على أداء مهام معينة تُحاكي وتُشابه آلية التفكير عند الإنسان؛ كالقدرة على التفكير، أو التعلّم من التجارب السابقة، أو غيرها من العمليات الأخرى التي تتطلب عمليات ذهنية. يوضح الشكل (3.2) أبرز مجالات الذكاء الاصطناعي.



الشكل (3.2): بعض مجالات الذكاء الاصطناعي

1-3-2 التعلم الآلي (Machine Learning):

التعلم الآلي هو برمجة أجهزة الكمبيوتر لتحسين مقاييس الأداء باستخدام بيانات نموذجية أو تجربة سابقة. لدينا نموذج محدد لبعض المعاملات، والتعلم هو تنفيذ برنامج كمبيوتر لتحسين معاملات النموذج باستخدام البيانات التعليمية أو الخبرة السابقة. قد يكون هذا النموذج تنبؤيًا لعمل تنبؤات في المستقبل، أو لاكتساب المعرفة من البيانات، أو كليهما [28]. يعني التعلم الآلي أن أجهزة الكمبيوتر تعمل أو تكيف إجراءاتها (سواء كانت تنبؤية أو تتحكم في روبوت) لجعلها أكثر دقة، حيث تنعكس الدقة في مقدار الإجراءات المحددة بشكل صحيح.

1-1-3-2 التعلم الخاضع للإشراف (Supervised Learning):

في طريقة التعلم الخاضع للإشراف، يتم تغذية مجموعة من عينات التدريب للخوارزمية بالإجابات الصحيحة (الأهداف) وتحاول الخوارزمية تعلم دالة بناءً على هذه البيانات والإجابات الصحيحة (يتعلم بمرور الوقت ويصبح أكثر دقة بمرور الوقت) لتكون قادرة على التنبؤ بدقة بالقيم المستهدفة للعينات الجديدة.

بمعنى آخر، الهدف هو تكييف النظام بطريقة تمكن النظام من التنبؤ بالمرجات الصحيحة للمدخلات الجديدة بناءً على ما تعلمه حتى الآن من بيانات التدريب. يسمى هذا النوع من طريقة التعلم أيضًا "التعلم من الأمثلة".

2-1-3-2 التعلم غير خاضع للإشراف (Unsupervised Learning):

في النهج غير خاضع للإشراف، لا يتم تقديم الإجابات الصحيحة للخوارزمية (لم يتم تصنيف البيانات)، ولكن بدلاً من ذلك تحاول الخوارزمية تحديد أوجه التشابه بين المدخلات بحيث يتم تجميع المدخلات التي لها ميزة مشتركة معًا.

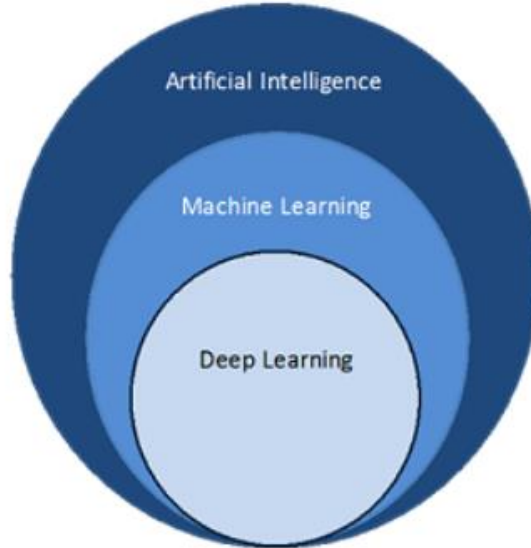
بمعنى آخر، لا يحتوي النظام على المخرجات المناسبة، لكنه يستكشف البيانات ويمكنه استخراج استنتاجات من مجموعة البيانات لوصف الهياكل المخفية للبيانات غير المسماة. التجميع وقواعد الارتباط وتقليل الأبعاد هي أمثلة على التعلم غير الخاضع للإشراف.

3-1-3-2 التعلم المعزز (Reinforcement Learning):

في التعلم المعزز، يحاول الوكيل حل مشكلة ما عن طريق التجربة والخطأ من خلال التفاعل مع بيئة طبيعتها غير معروفة للوكيل. يمكن للوكيل تغيير حالة البيئة من خلال أفعاله أثناء تلقي تعليقات فورية من البيئة. هدف الوكيل هو حل المشكلة من خلال إيجاد سلسلة الإجراءات المثلى. على الرغم من أن التعلم المعزز هو أحد مجالات التعلم الآلي، إلا أنه يختلف اختلافاً جوهرياً عن أساليب التعلم الآلي القياسية (الإشرافي وغير الإشرافي) من نواح كثيرة. أولاً، لا يعتمد التعلم المعزز على تعلم البيانات. بدلاً من ذلك، يتعلم الوكيل من خبرته المكتسبة أثناء التفاعل مع البيئة ولا يعتمد على المشرف. ثانياً، يركز التعلم المعزز على إيجاد السياسة المثلى بدلاً من تحليل البيانات.

2-3-2 التعلم العميق (Deep learning):

التعلم العميق هو مجموعة فرعية من التعلم الآلي الذي يعلم الآلات القيام بالأشياء التي يولد بها البشر بشكل طبيعي مثل: التعلم من خلال الأنماط. على الرغم من أن هذه التقنية غالباً ما تُعتبر مجموعة من الخوارزميات التي "تحاكي الدماغ"، فإن الوصف الأكثر ملاءمة هو مجموعة من الخوارزميات التي "تتعلم من خلال الطبقات". بمعنى آخر، يتضمن التعلم من خلال الطبقات التي تمكن الخوارزمية من إنشاء تسلسل هرمي للمفاهيم المعقدة من المفاهيم الأبسط كما يبين الشكل (2.4).



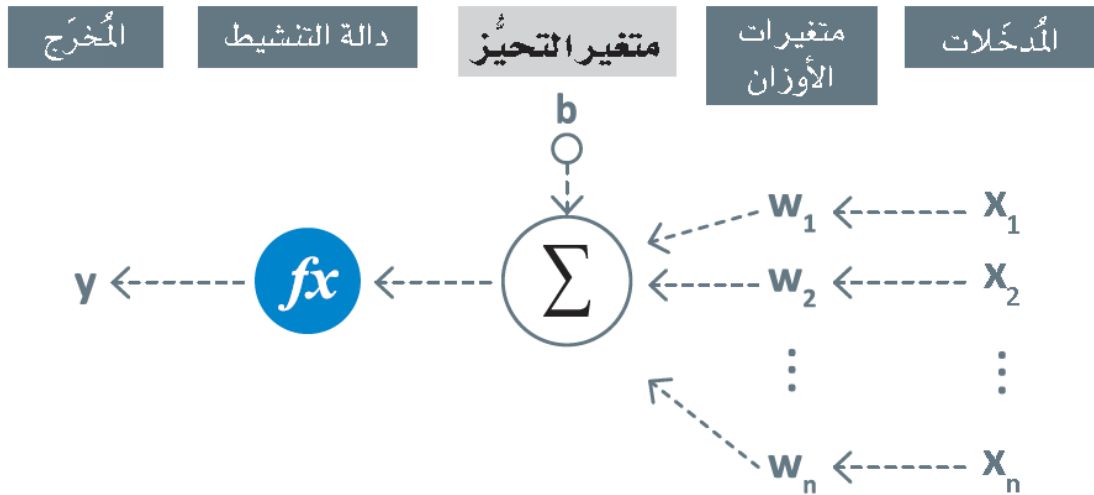
الشكل (4.2) التعلم العميق

يصف التعلم العميق الخوارزميات التي تحلل البيانات بهيكل منطقي، على غرار الطريقة التي يستنتجها الإنسان. نلاحظ أن هذا يمكن أن يحدث من خلال التعلم الخاضع للإشراف وغير الخاضع للإشراف. لتحقيق هذا الهدف، تستخدم تطبيقات التعلم العميق بُنية ذات طبقات (هرمية) من الخوارزميات تسمى "الشبكات العصبية الاصطناعية".

1-2-3-2 الشبكة العصبية:

الخلايا العصبية الاصطناعية:

الخلايا العصبية الاصطناعية (العصبونات) هي اللبنة الرئيسية للشبكات العصبية الاصطناعية. تصف الخلايا العصبية الاصطناعية نموذجاً رياضياً بسيطاً مستوحى من الخلايا العصبية للدماغ، تتمثل الوظيفة الأساسية للخلايا العصبية الاصطناعية في تلقي مدخلات متعددة x_1, \dots, x_n وحساب مجموع الأوزان w_1, \dots, w_n لهذه المدخلات باستخدام الأوزان w_1, \dots, w_n مجموعة الأوزان z هو تحويل خطي لمدخلات الخلايا العصبية. بالإضافة إلى ذلك، يضاف التحيز b إلى مجموع الأوزان للمدخلات ويتم تمرير النتيجة من خلال دالة التنشيط φ ، مما ينتج عنه إخراج نهائي \hat{y} كما يبين الشكل (5-2).



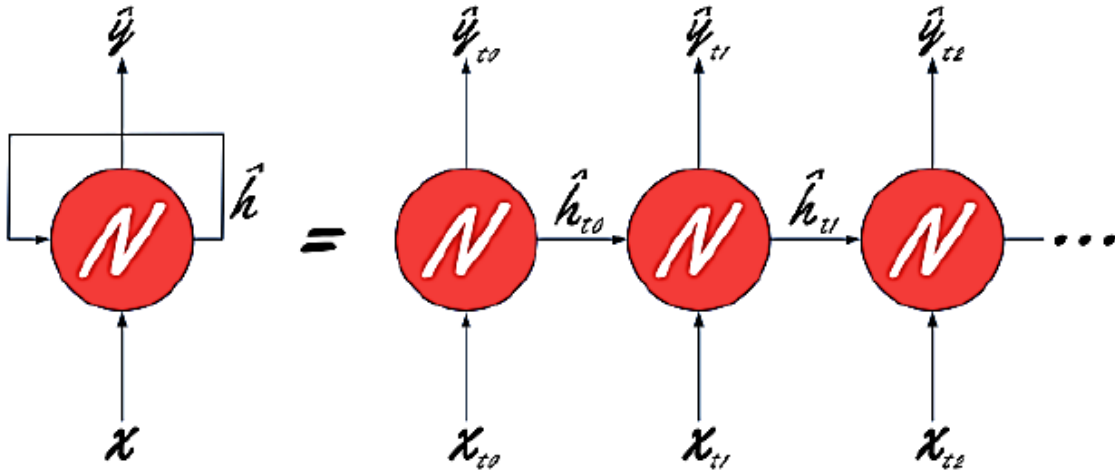
الشكل (5-2): الخلية العصبية الصناعية.

الشبكة العصبية المتكررة:

الشبكات العصبية المتكررة (RNN) هي نوع من الشبكات العصبية الاصطناعية المصممة لاكتشاف الأنماط في تسلسل البيانات، مثل النص والكتابة اليدوية والكلمات المنطوقة وبيانات السلاسل الزمنية وأسواق الأسهم وما إلى ذلك. الفكرة من وراء هذه الشبكات العصبية هي أنها تسمح للخلايا بالتعلم

من الخلايا المرتبطة سابقاً. يمكن القول إن هذه الخلايا لها "ذاكرة" بطريقة ما. ومن ثم، تبني معرفة أكثر تعقيداً من بيانات الإدخال.

تأخذ الشبكات العصبية المتكررة تسلسلاً كمدخلات وتقيم الشبكة العصبية لكل خطوة زمنية. يمكن اعتبار هذه الشبكات على أنها شبكة عصبية لها حلقة تسمح لها بالحفاظ على الحالة. عند التقييم، تفتح الحلقة من خلال الخطوات الزمنية للتسلسل كما هو موضح في الشكل المقابل. هذه الحلقات أو الروابط المتكررة هي سبب تسمية هذه الشبكات بالشبكات العصبية المتكررة.



الشكل (6-2): الشبكة العصبية المتكررة

4-2 كشف الهجمات السيبرانية المعتمد على الذكاء الاصطناعي:

في الآونة الأخيرة، أثرت تقنيات الذكاء الاصطناعي (AI) بشكل كبير على تطوير وتطبيق الأمن السيبراني في مختلف الأنظمة والشبكات. هناك العديد من تقنيات الذكاء الاصطناعي ولكن الاستخدام الأكثر شيوعاً للذكاء الاصطناعي في الأمن السيبراني يأتي من منظور التعلم الآلي (Machine Learning). خوارزميات التعلم الآلي هي فرع من مجال الذكاء الاصطناعي الأوسع نطاقاً، تستخدم هذه الخوارزميات النماذج الإحصائية لتطوير التنبؤات. لذلك يمكن أن تكون المخرجات المتوقعة هي الأكثر دقة لنظام اكتشاف الهجمات. حيث أنه بناءً على صعوبة التحكم بشكل دقيق في مكان ووقت وكيفية حدوث الهجمات السيبرانية، وعدم وجود ضمان للوقاية الكاملة، يبدو أن الكشف المبكر هو أفضل نهج لمواجهة المخاطر التي قد لا يُمكن إصلاح أضرارها. وتكون النتيجة المثالية عندما تستفيد

المؤسسات إما من الحلول الموجودة بالفعل أو تطور حلولها الخاصة لكشف الهجمات السيبرانية حيث يتم تقليل الحاجة إلى تدخل الإنسان.

تم تنفيذ تقنيات مختلفة للتعلم الآلي (Machine Learning) للكشف عن اختراق النظام والشبكة بهدف تحسين معدلات الكشف والحفاظ على قاعدة معارفهم شاملة وحديثة، ومع التحسينات في أنظمة الأمن السيبراني، فإن القدرة على التكيف هي السمة الأساسية للأنظمة الأمنية السيبرانية [29].

كذلك، يمكن لخوارزميات التعلم العميق التعلّم من مجموعات بيانات البريد الإلكتروني المصنفة لبناء نماذج قادرة على التعرف على الأنماط التي تشير إلى رسائل البريد الإلكتروني المزعجة، حيث أن هناك العديد من المميزات لاستخدام التعلم العميق للكشف عن رسائل البريد الإلكتروني المزعجة، فهو يسمح بالترشيح الفعّال وفصل رسائل البريد الإلكتروني المشروعة عن البريد المزعج، مما يقلل من الوقت والجهد الذي يبذله المستخدمون في الفرز اليدوي لصندوق الوارد الخاص بهم. بعد ذلك، يمكن استخدام هذه النماذج لتصنيف رسائل البريد الإلكتروني الجديدة غير المرئية تلقائياً. وذلك من خلال تحليل خصائص البريد الإلكتروني المختلفة مثل معلومات المرسل و سطر الموضوع والمحتوى وروابط URL المضمنة، يمكن لخوارزميات التعلم العميق تحديد خصائص البريد المزعج وتقديم تنبؤات دقيقة.

هناك العديد من تقنيات التعلم الآلي المستخدمة بشكل شائع للكشف عن الهجمات السيبرانية. وتشمل هذه التقنيات طريقة (Support Vector Machine) وأشجار القرار (Decision Trees) والغابات العشوائية (Random forest) وخوارزمية (Naive Bayes) والشبكات العصبية (Neural Networks) وخوارزميات التعلم العميق (Deep learning) .

يمكن تدريب هذه الخوارزميات على مجموعات البيانات المصنفة، مما يسمح لها بتعلم الأنماط والعلاقات الأساسية بين الهجمات الخبيثة وحركات المرور الطبيعية.

الفصل الثالث:

منهجية العمل

3-1 لمحة عامة:

تُعتبر منهجية البحث أحد العناصر الأساسية في أي دراسة علمية، فهي تعبّر عن الخطوات والطرق المتبعة لتحقيق الأهداف المحددة وفحص الفرضيات المطروحة. يهدف هذا الفصل في بحثنا إلى توضيح الخطوات والتقنيات التي تم اتباعها في تنفيذ منهجية البحث حول كشف أبرز الهجمات السيبرانية إضافة إلى كشف الهجمات من نوع رسائل البريد الإلكتروني التصيدي على مواقع الويب باستخدام الذكاء الاصطناعي. تتألف منهجية البحث في هذه الدراسة من سلسلة من الخطوات التي تبدأ بالحصول على البيانات وتنتهي بتقييم النموذج المقترح. يتضمن الفصل شرحاً لكل خطوة من الخطوات المتبعة، بما في ذلك تحليل البيانات، ومعالجتها، وتحليل النصوص، وتطبيق تقنيات الذكاء الاصطناعي، وتدريب وتقييم النموذج. تهدف هذه الخطوات إلى تحقيق أهداف البحث بشكل شامل ودقيق. حيث تُوفر هذه المنهجية الإرشادات والتوجيهات اللازمة لتحقيق الأهداف البحثية بطريقة منظمة وفعّالة، مما يساهم في جودة وموثوقية النتائج التي تم الحصول عليها وتحقيق الفوائد العلمية المرجوة من البحث. وبما أن مجموعة بيانات (Dataset) الخاصة برسائل البريد الإلكتروني التصيدي منفصلة ومختلفة عن مجموعة بيانات أبرز الهجمات السيبرانية، لذلك سيكون لدينا منهجيتين مختلفتين، ولكن يبقى الهدف نفسه وهو إيجاد نماذج الذكاء الاصطناعي التي تقوم بعملية كشف مختلف الهجمات بدقة عالية مما يؤمن الحماية لمواقع الويب الإلكترونية المعرضة بشكل دائم للهجمات السيبرانية.

3-2 أدوات البحث:

3-2-1 لغة البايثون (Python):

بايثون هي لغة برمجة عالية المستوى موجهة للكائنات وتستخدم لمجموعة واسعة من المشكلات ذات النطاق والتعقيد المتفاوتين. على عكس العديد من اللغات المماثلة، فهي تعتبر من أكثر اللغات شعبية وانتشاراً، وأهم الميزات التي تتميز بها هذه اللغة هي:

- البساطة: لغة البرمجة بايثون هي واحدة من أسهل اللغات للبدء بها، مع العلم أن هذه البساطة لا تحد من الميزات التي يحتاجها أي باحث أو خبير وفي مجالات متنوعة.
- المكتبات والأطر: تمتلك لغة البرمجة بايثون المئات من المكتبات والأطر المختلفة التي تساعد بشكل كبير في عملية التطوير وتوفر الكثير من الوقت.
- مجتمع هائل: أحد أسباب شهرة لغة البرمجة بايثون هو أنها تضم مجتمعًا كبيرًا من المهندسين والعلماء.
- أهميتها في التعلم العميق: تحتوي لغة البرمجة بايثون على العديد من الحزم مثل keras و Tensorflow و PyTorch التي تساعد على تنفيذ تطوير خوارزميات التعلم العميق.

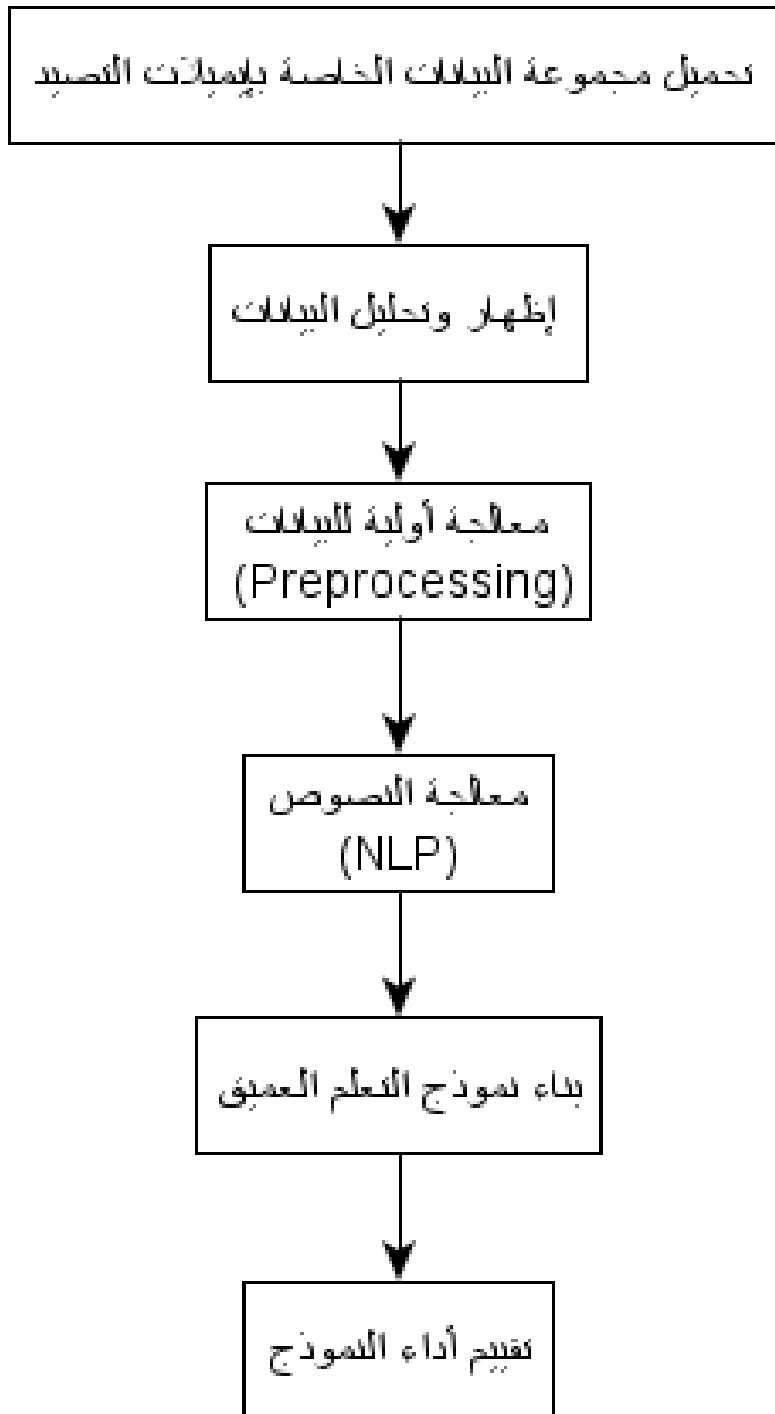
3-2-2 نوتبوك جوبيتر Jupyter Notebook:

نوتبوك جوبيتر هو أداة قوية بشكل لا يصدق لتطوير وتقديم مشاريع علم البيانات التفاعلية التي يمكن أن تتضمن نصًا أو صورة أو صوتًا أو فيديو بالإضافة إلى تنفيذ التعليمات البرمجية. حيث يجمع النوتبوك Notebook بين التعليمات البرمجية والمخرجات مع الرسوم التوضيحية والنص السردي والمعادلات الرياضية والوسائط الأخرى في مستند واحد. بمعنى آخر، يمكن من خلال نوتبوك جوبيتر تنفيذ التعليمات البرمجية وعرض الإخراج وإضافة الأوصاف والصيغ والرسوم التخطيطية لجعل العمل أكثر وضوحًا وقابلية للفهم وقابلية للتكرار والمشاركة.

3-3 المنهجية المقترحة لكشف رسائل البريد الإلكتروني التصيدي:

أصبح اختراق البريد الإلكتروني (تصيد البيانات) تهديدًا كبيرًا للأفراد والمؤسسات على مستوى العالم. تهدف رسائل البريد الإلكتروني المخادعة هذه إلى تضليل المستلمين للكشف عن معلومات حساسة أو اتخاذ إجراءات ضارة.

يعتبر اكتشاف رسائل البريد الإلكتروني الاحتيالية ومنعها أمرًا ضروريًا لحماية الأمن الشخصي والمالي. في السنوات الأخيرة، ظهرت تقنيات التعلم الآلي كنهج واعد لمكافحة هذا الخطر المتزايد. يتمثل المخطط المقترح لمراحل العمل الأساسية في سلسلة من الخطوات المنظمة والمتسلسلة لمنهجية البحث المقترحة لهذا النوع من الهجمات كما يوضح الشكل (3-1).



الشكل (1-3) منهجية العمل المقترحة لكشف رسائل البريد التصيدي

3-3-1 تحميل البيانات ذات الصلة:

0	18.6k	empty 3%	Safe Email 61%
		[null] 0%	Phishing Email 39%
		Other (18101) 97%	
3		Hello I am your hot lil horny toy. I am the one you dream About, I am a very open minded pe...	Phishing Email
4		software at incredibly low prices (86 % lower) . drapery seventeen term represent any sing . feet ...	Phishing Email
5		global risk management operations sally congratulations on your new role . if you were not already a...	Safe Email
6		On Sun, Aug 11, 2002 at 11:17:47AM +0100,	Safe Email

الشكل (2-3): مجموعة البيانات الخاصة برسائل البريد الإلكتروني السليمة والتصيدية

مجموعة البيانات المستخدمة للكشف عن رسائل البريد الإلكتروني السليمة والتصيدية هي المجموعة "Phishing Email Detection" [30]، تم تحميلها من موقع Kaggle (إحدى الشركات التابعة لشركة جوجل، وهي عبارة عن مجتمع عبر الإنترنت لعلماء البيانات ومهندسي التعلم الآلي).

تم التأكد من أن البيانات في صيغة مناسبة للاستخدام مع أدوات التعلم الآلي كما يوضح الشكل (3-2). تحدد مجموعة البيانات نص جسم البريد الإلكتروني ونوع رسائل البريد الإلكتروني التي يمكن استخدامها للكشف عن رسائل البريد الإلكتروني الاحتيالية من خلال تحليل شامل لنص البريد الإلكتروني وتصنيفها باستخدام التعلم الآلي.

3-3-2 مرحلة إظهار وتحليل البيانات:

مجموعة البيانات التي تم تحميلها هي عبارة عن مجموعة من رسائل بريد إلكتروني تم تصنيفها على أنها "تصيد احتيالي" (Phishing Email) أو "شرعية - آمنة" (Safe Email)، كما يبين الشكل (3-3).

Unnamed: 0	Email Text	Email Type
0	re : 6 . 1100 , disc : uniformitarianism , re ...	Safe Email
1	the other side of * galicismo * * galicismo *...	Safe Email
2	re : equistar deal tickets are you still avail...	Safe Email
3	\nHello I am your hot lil horny toy.\n I am...	Phishing Email
4	software at incredibly low prices (86 % lower...	Phishing Email

الشكل (3-3): إظهار أول خمس أسطر من مجموعة البيانات

- تم حساب عدد العينات الكلي لمدخلات مجموعة البيانات المستخدمة كما في الشكل (4-3).

Number of samples in the dataset: 18650

الشكل (4-3): عدد العينات الكلي في مجموعة البيانات المستخدمة

- تم حساب عدد العينات لكل صنف من أصناف مجموعة البيانات وهي عبارة عن صنفين أساسيين: الصنف الأول هو رسائل البريد الإلكتروني الآمنة وعدد عيناتها 11322، والصنف الثاني هو رسائل التصيد وعددها 7328، كما يبين الشكل (5-3).

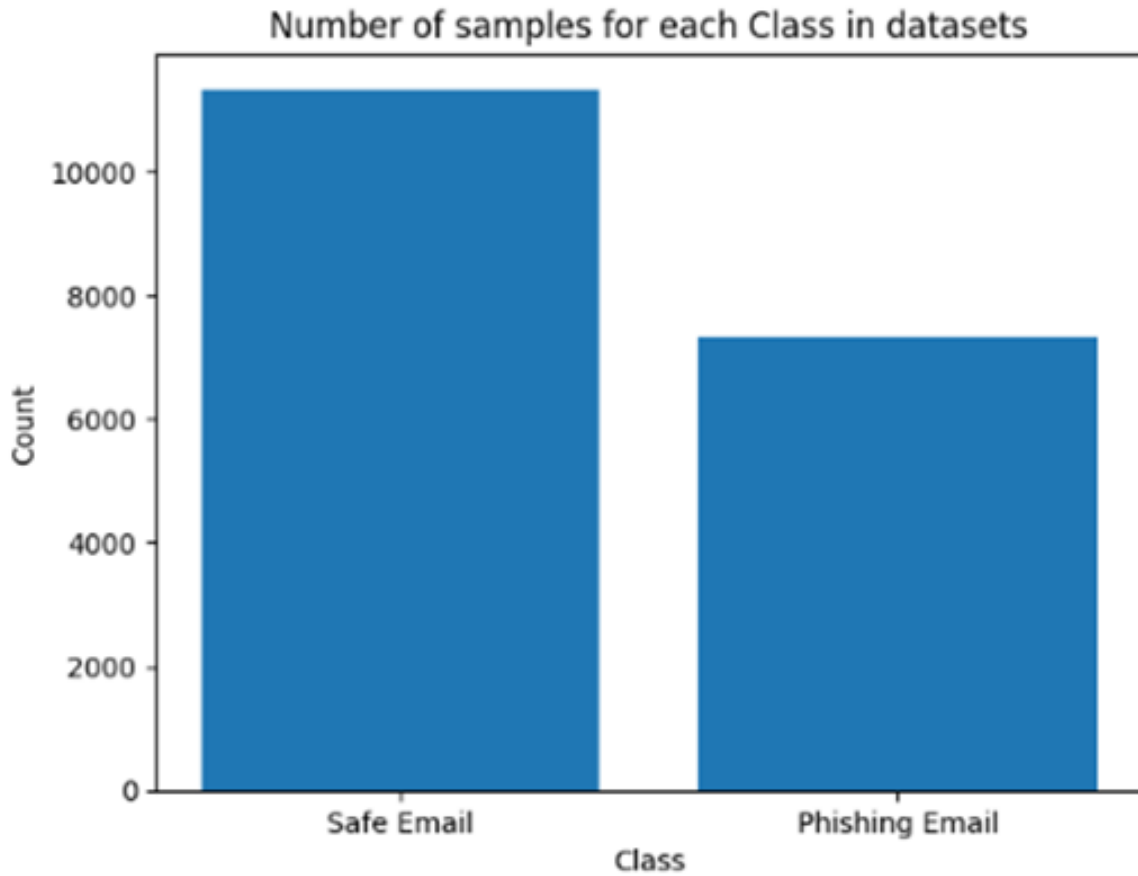
```

Classes in the dataset:
Safe Email      11322
Phishing Email  7328

```

الشكل (3-5): الأصناف في مجموعة البيانات

- تم استعراض عدد العينات لكل صنف على شكل مخطط كما في الشكل (3-6).



الشكل (3-6): مخطط للأصناف في مجموعة البيانات

- تم إظهار نوع البيانات المستخدمة، ليتبين كما يوضح الشكل (3-7) أن العمود الأول يحتوي

على أرقام بينما يحتوي العمودين الثاني والثالث على كائنات.

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 18650 entries, 0 to 18649
Data columns (total 3 columns):
#   Column          Non-Null Count  Dtype
---  ---
0   Unnamed: 0      18650 non-null  int64
1   Email Text      18634 non-null  object
2   Email Type      18650 non-null  object
dtypes: int64(1), object(2)
memory usage: 437.2+ KB

```

الشكل (7-3): نوع البيانات في مجموعة البيانات

3-3-3 المعالجة الأولية للبيانات:

تشير المعالجة الأولية للبيانات إلى الخطوات الأولية التي يتم اتخاذها لتنظيف وتنظيم البيانات الواردة قبل تحليلها أو استخدامها، وتتضمن ما يلي:

- إزالة الأعمدة غير المرغوبة: يوجد في مجموعة البيانات عمود مكرر يحتوي على ترقيم تسلسلي للأسطر، يتم حذف هذا العمود فتصبح مجموعة البيانات بعد الحذف كما في الشكل (8-3)، وكذلك يتم إضافة عمود يشير إلى رسائل البريد التي تحتوي على روابط ضمنها، رسالة البريد التي تحتوي على رابط يتم وضع قيمة 1 مقابلة له في هذا العمود وإلا تكون القيمة صفر.

	Email Text	Email Type	URL Existed
0	re : 6 . 1100 , disc : uniformitarianism , re ...	Safe Email	0
1	the other side of * galicismos * * galicismo *...	Safe Email	0
2	re : equistar deal tickets are you still avail...	Safe Email	0
3	\nHello I am your hot lil horny toy.\n I am...	Phishing Email	1
4	software at incredibly low prices (86 % lower...	Phishing Email	0

الشكل (8-3): مجموعة البيانات بعد حذف عمود غير ضروري

- إزالة جميع البيانات الناقصة أو غير المتسقة: في البداية يتم استعراض المدخلات التي تحتوي على قيم فارغة كما في الشكل (9-3).

```

Unnamed: 0      0
Email Text     16
Email Type      0
dtype: int64

```

الشكل (3-9): البيانات الفارغة في مجموعة البيانات

- بعد إزالة القيم الفارغة من مجموعة البيانات نحصل على مجموعة بيانات نظيفة كما في الشكل (3-10).

```

Unnamed: 0      0
Email Text      0
Email Type      0
dtype: int64

```

الشكل (3-10): التأكد من إزالة البيانات الفارغة من مجموعة البيانات

- فتصبح قيم أصناف مجموعة البيانات بعد عمليات المعالجة الأولية كما في الشكل (3-11).

```

Classes in the dataset after removing Null values:
Safe Email      11322
Phishing Email  7312
Name: Email Type, dtype: int64

```

الشكل (3-11): قيم أصناف مجموعة البيانات بعد عمليات المعالجة الأولية

- إزالة علامات الترقيم والأحرف غير المرغوب فيها وكلمات التوقف التي تعتبر خطوة مهمة جداً في المعالجة المسبقة للنصوص حيث يتم استبدال كل منها بفاغ ومن ثم إعادة النص بشكل كامل. من أجل كل نص في كل إيميل يتم إزالة الفراغات والأسطر الفارغة ومن أجل الدمج أكثر يتم إزالة كل المحارف ما عدا المحارف من a to z (أحرف صغيرة وكبيرة) ومن 0 to 9. وأخيراً تقوم بدمج الكلمات ووضع بينها فراغات من أجل الحصول على جملة مع إزالة كلمات التوقف.

3-3-4 مرحلة تحليل النصوص باستخدام معالجة اللغات الطبيعية NLP:

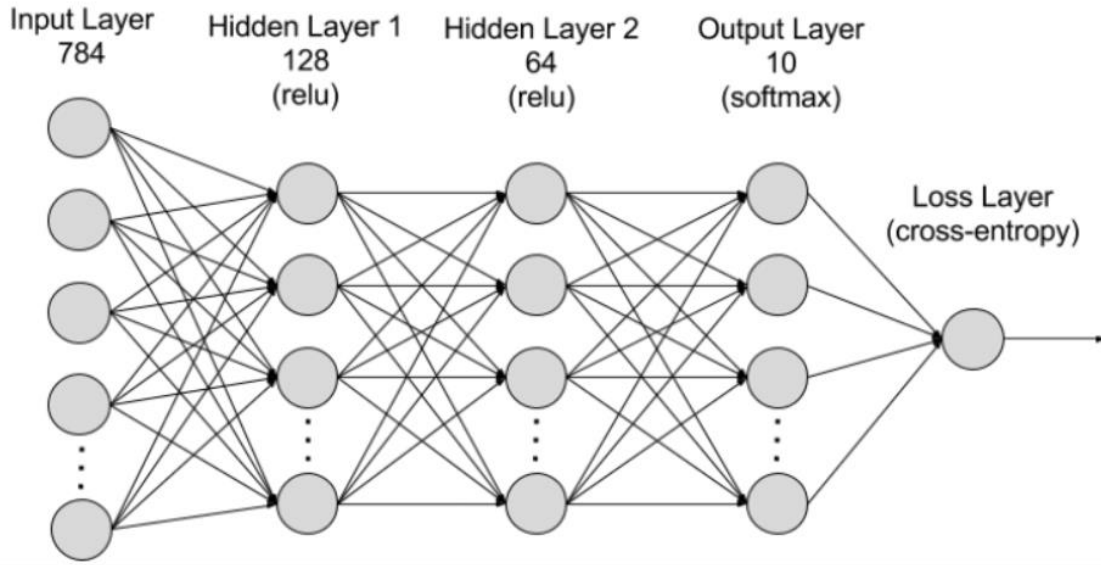
مرحلة تحليل النصوص باستخدام معالجة اللغات الطبيعية (NLP) تتضمن عدة خطوات أساسية تهدف إلى فهم واستخلاص المعلومات من النصوص بشكل آلي:

- حددنا (Tokenizer) يقوم بتحويل الجمل إلى كلمات وطبقنا Sequence الذي يقوم بتحويل كل كلمة إلى رقم، حيث يعطي لكل كلمة index ويستبدلها بكل النصوص، وهذه العملية هي نوع من أنواع معالجة اللغات الطبيعية من أجل تحويل الكلام إلى أرقام.

- عرفنا (pad_sequences) يقوم بتحويل كامل الجمل إلى جمل بطول 200، فمثلاً إذا وُجدت جمل بطول 100 محرف يقوم بعملية إكمال لها إلى الطول 200 محرف باستخدام رقم محدد للإكمال لأنه يجب على الدخل أن يكون موحد.
- بعد ذلك تمت عملية ترميز للهدف، لأن الهدف يكون إما Safe or Phishing فيجب ترميزه أيضاً بأرقام، فيتم إعطاء 0 لرسائل البريد من نوع تصيد و 1 لرسائل البريد الآمنة.
- أخيراً، تم تقسيم مجموعة البيانات إلى بيانات تدريب وبيانات اختبار، بحيث يكون لدينا 80% لبيانات التدريب و 20% لبيانات الاختبار.

3-3-5 بناء نموذج الشبكة العصبية العميقة (Deep Neural Network):

الشبكة العصبية العميقة هي نوع من النماذج الرياضية المستوحاة من الدماغ البشري تُستخدم لتعلم الأنماط المعقدة في البيانات. تتكون الشبكات العصبية العميقة من طبقات متعددة (طبقة إدخال وطبقات مخفية وطبقة إخراج) كما يوضح الشكل (3-12)، كل منها قادرة على التعلم واستخراج الميزات من البيانات، يسمى نهج الشبكة العصبية متعدد الطبقات بالتعلم العميق.



الشكل (3-12): بنية الشبكة العصبية العميقة

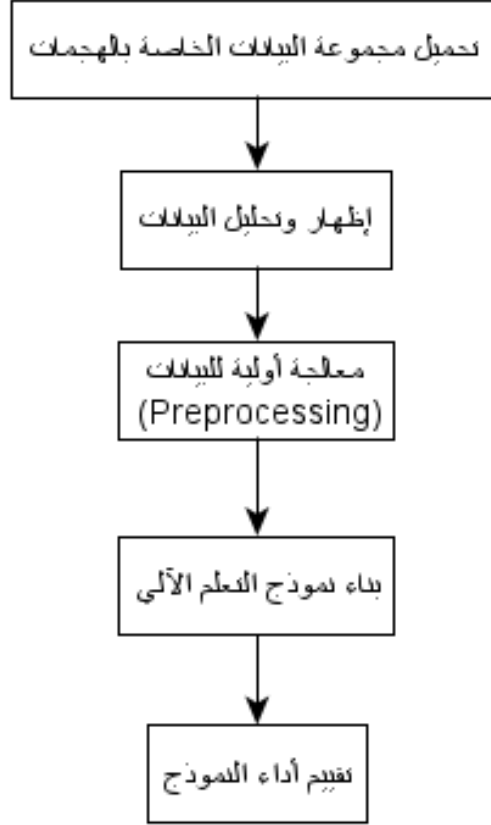
3-3-1 الشبكات العصبية التكرارية (RNNs):

هي نوع من الشبكات العصبية العميقة في معالجة البيانات التسلسلية مثل النصوص والكلام والترجمة والموسيقى. تتميز RNN بالقدرة على الاحتفاظ بذاكرة داخلية تمكنها من فهم السياق والتنبع عبر الزمن. تكون البيانات المدخلة ذات بعد واحد على شكل متجهات أو سلاسل رقمية (عدد العناصر في السلسلة). ومع ذلك، فإن RNNs لها بعض القيود، بما في ذلك صعوبة تعلم التبعية بعيدة المدى بسبب مشكلة التدرج المتلاشي. تمت معالجة هذه المشكلة جزئياً بواسطة بنيات RNN الأكثر تقدماً، مثل الوحدات ذات البوابات المتكررة (GRU)، والتي تقدم آليات متخصصة لمساعدة الشبكة على التعلم والاحتفاظ بالتبعية طويلة المدى بشكل أكثر فعالية. تم تعريف الشبكة العصبية RNN من نوع GRU مع Bidirectional. يشير مصطلح Bidirectional إلى أن الشبكة تعمل في كلا الاتجاهين، أي أنها تحاول فهم السياق والترتيب في البيانات من اليمين إلى اليسار ومن اليسار إلى اليمين. هذا يساعد في فهم السياق بشكل أفضل خاصة في المجالات التي تعتمد على السياق مثل تحليل النصوص.

بالجمع بين هذه العناصر، تصبح الشبكة العصبية العميقة من نوع GRU مع Bidirectional قادرة على فهم السياق والترتيب في البيانات بشكل أفضل من خلال عملية ربط الخطوات السابقة مع اللاحقة من أجل تحقيق نوع أكثر من الترابط بين الجمل فتكون الكلمة الأولى لها علاقة بالكلمة الثانية والكلمة الثالثة والكلمة الرابعة والخامسة من أجل أن نأخذ المعنى بشكل كامل.

3-4 المنهجية المقترحة لكشف الهجمات على مواقع الويب:

تتعرض مواقع الويب باستمرار لمجموعة متنوعة من الهجمات، تتراوح من اختراقات البرامج الضارة إلى هجمات SQL Injection. تشكل هذه الهجمات تهديداً خطيراً لسلامة البيانات وسرية المعلومات، مما قد يؤدي إلى خسائر مالية وسمعة سيئة. تهدف هذه الدراسة التحليلية إلى فهم أنواع هجمات الويب المختلفة وتأثيرها على المواقع الإلكترونية، واستخدام تقنيات التعلم الآلي وتقييم فعاليتها في الكشف عن الهجمات الشهيرة على مواقع الويب. المنهجية المقترحة لهذه الدراسة موضحة في الشكل (3-13).



الشكل (3-13): المنهجية المقترحة لكشف الهجمات على مواقع الويب

3-4-1 تحميل مجموعة البيانات الخاصة بالهجمات على مواقع الويب:

مجموعة البيانات المستخدمة هي (IDS 2018 Intrusion CSVs (CSE-CIC-IDS2018) وتم تحميلها من موقع kaggle . تم إنشاء هذه المجموعة في الأصل بواسطة جامعة نيو برونزويك لتحليل بيانات هجمات منع الخدمة الموزعة (DDoS). تستند مجموعة البيانات نفسها على سجلات خوادم الجامعة، والتي عثرت على العديد من هجمات منع الخدمة (DoS) طوال الفترة المتاحة للجمهور، عند قراءة مواصفات هذه المجموعة من البيانات، نلاحظ أن عمود "العلامة" (Label) هو بلا شك أهم جزء من البيانات، حيث أنه يحدد ما إذا كانت الحزم المرسلّة ضارة أم لا. في المجموع، هناك ثمانون عمودًا داخل هذه المجموعة، يتوافق كل منها مع إدخال في نظام تسجيل IDS الذي تمتلكه جامعة نيو برونزويك. نظرًا لأن نظامهم يصنف حركة المرور على أنها حركة أمامية وخلفية، فهناك أعمدة لكليهما. يتم سرد أهم الأعمدة ضمن هذه المجموعة أدناه.

- منفذ الوجهة (Dst Port)
- بروتوكول (Protocol)
- مدة التدفق (Flow Duration)
- إجمالي حزم التوجيه الأمامية (Tot Fwd Pkts)
- إجمالي حزم التوجيه الخلفية (Tot Bwd Pkts)
- الوسم (Label)

3-4-2 قراءة وتحليل مجموعة البيانات:

33	Bwd PSH Flags	999984	non-null	object
34	Fwd URG Flags	999984	non-null	object
35	Bwd URG Flags	999984	non-null	object
36	Fwd Header Len	999984	non-null	object
37	Bwd Header Len	999984	non-null	object
38	Fwd Pkts/s	999984	non-null	object
39	Bwd Pkts/s	999984	non-null	object
40	Pkt Len Min	999984	non-null	object
41	Pkt Len Max	999984	non-null	object
42	Pkt Len Mean	999984	non-null	object
43	Pkt Len Std	999984	non-null	object
44	Pkt Len Var	999984	non-null	object
45	FIN Flag Cnt	999984	non-null	object
46	SYN Flag Cnt	999984	non-null	object
47	RST Flag Cnt	999984	non-null	object
48	PSH Flag Cnt	999984	non-null	object
49	ACK Flag Cnt	999984	non-null	object
50	URG Flag Cnt	999984	non-null	object
51	CWE Flag Count	999984	non-null	object
52	ECE Flag Cnt	999984	non-null	object
53	Down/Up Ratio	999984	non-null	object
54	Pkt Size Avg	999984	non-null	object
55	Fwd Seg Size Avg	999984	non-null	object
56	Bwd Seg Size Avg	999984	non-null	object
57	Fwd Byts/b Avg	999984	non-null	object
58	Fwd Pkts/b Avg	999984	non-null	object
59	Fwd Blk Rate Avg	999984	non-null	object
60	Bwd Byts/b Avg	999984	non-null	object
61	Bwd Pkts/b Avg	999984	non-null	object
62	Bwd Blk Rate Avg	999984	non-null	object
63	Subflow Fwd Pkts	999984	non-null	object
64	Subflow Fwd Byts	999984	non-null	object
65	Subflow Bwd Pkts	999984	non-null	object
66	Subflow Bwd Byts	999984	non-null	object
67	Init Fwd Win Byts	999984	non-null	object
68	Init Bwd Win Byts	999984	non-null	object
69	Fwd Act Data Pkts	999984	non-null	object
70	Fwd Seg Size Min	999984	non-null	object
71	Active Mean	999984	non-null	object
72	Active Std	999984	non-null	object
73	Active Max	999984	non-null	object

الشكل (3-14): جانب من خصائص مجموعة البيانات

تحتوي مجموعة البيانات على 10 ملفات تم تسجيلها خلال فترات زمنية مختلفة، قمنا بأخذ 100000 سجل من كل ملف من هذه الملفات ليصبح في النهاية عدد مدخلات مجموعة البيانات المستخدمة حوالي مليون مدخل، بينما يكون عدد الخصائص 80 خاصية متمثلة بالأعمدة كما يوضح الشكل (3-14) جانب من خصائص (Features) مجموعة البيانات.

الهجمات الشهيرة على مواقع الويب التي تحتويها مجموعة البيانات المستخدمة هي:

- DDoS attacks-LOIC-HTTP
- FTP-BruteForce
- DDOS attack-HOIC
- DDOS attack-LOIC-UDP
- DoS attacks-SlowHTTPTest
- DoS attacks-GoldenEye
- DoS attacks-Slowloris
- DoS attacks-Hulk
- Bot
- Brute Force –Web
- SQL Injection
- Brute Force –XSS

توزيع العينات لكل صنف من أصناف الهجمات موضح في الشكل (3-15).

Label	
Benign	465279
DDoS attacks-LOIC-HTTP	99918
FTP-BruteForce	99882
DDOS attack-HOIC	96064
DoS attacks-SlowHTTPTest	91434
Bot	80182
DoS attacks-GoldenEye	41508
DoS attacks-Slowloris	10990
DoS attacks-Hulk	8300
DDOS attack-LOIC-UDP	1730
Brute Force -Web	611
Brute Force -XSS	230
SQL Injection	87

الشكل (3-15): توزيع العينات على أصناف الهجمات المستخدمة

3-4-3 المعالجة الأولية للبيانات:

تتضمن الخطوات المتعلقة بالتحضير (Preprocessing) لمجموعة البيانات ما يلي:

- إزالة العمود (Timestamp) غير الضروري: في هذه الخطوة، قمنا بإزالة العمود "Timestamp" كونه غير مهم حيث يحتوي على توقيت حصول الهجوم.
- تحويل كل عمود إلى نوع البيانات الرقمي (int): قمنا بتحويل كل عمود في مجموعة البيانات إلى نوع البيانات الرقمي (integer)، بحيث يكون لدينا قيم رقمية قابلة للتحليل والاستخدام في التدريب.
- استبدال القيم المجهولة بالقيم الفارغة (null): قمنا بتعويض أي قيم مجهولة أو غير معروفة في مجموعة البيانات بقيم فارغة (null)، لتسهيل معالجة البيانات في الخطوات اللاحقة.
- حذف جميع القيم الفارغة (null): قمنا بحذف أي صفوف في مجموعة البيانات التي تحتوي على قيم فارغة (null)، لتجنب أي تأثير سلبي على عمليات التحليل أو التدريب اللاحقة.
- تقسيم عمود العلامات: قمنا بتقسيم مجموعة البيانات إلى مجموعتين: البيانات الخاصة بخصائص مجموعة البيانات (features) والتي سستخدم للتنبؤ، والبيانات الخاصة بنتيجة التصنيف (label) الذي سنحاول التنبؤ به.

وأخيراً يتم تقسيم مجموعة البيانات إلى قسم تدريب بنسبة 80% من مجموع البيانات، وقسم آخر للاختبار 20% وهو يمثل البيانات الجديدة التي سوف يتم استخدامها لاختبار النموذج وتقييم أدائه.

3-4-4 نموذج التعلم الآلي (شجرة القرار):

شجرة القرار واحدة من أكثر خوارزميات التعلم الآلي شيوعاً، نظراً لعملياتها البسيطة جداً. إن أشجار القرار تحاكي حرفياً الطريقة التي نقوم بها نحن البشر على أساس يومي في تحليل القرار، حيث يمكن استخدام أشجار القرار لاتخاذ القرارات بشكل موضوعي وصريح -وكما يوحي الاسم- يتم استخدام نموذج يشبه الشجرة للوصول إلى القرار النهائي.

عادةً ما تتكون شجرة القرار من عقدة جذر والعديد من العقد الداخلية والعديد من العقد الطرفية. تتوافق العقد الطرفية مع نتائج القرار، وتتوافق كل عقدة أخرى مع اختبار الخاصية، يتم تقسيم العينات في كل عقدة إلى عقد فرعية وفقاً لنتائج التفرع. كل مسار من العقدة الجذرية إلى العقدة الطرفية هو تسلسل قرار. تحاول أشجار القرار تفريع مجموعة البيانات بطريقة تجعل البيانات في كل مجموعة متشابهة قدر

الإمكان، بينما تختلف البيانات في مجموعة واحدة قدر الإمكان عن البيانات في المجموعات الأخرى، الهدف هو إنتاج شجرة يمكنها تعميم العينات.

يتبع بناء أشجار القرار استراتيجية (فرق تسد)، بعبارة أبسط، تبدأ خوارزمية شجرة القرار بمجموعة بيانات التدريب في عقدة الجذر وتقسّم البيانات بشكل متكرر إلى عقد ذات مستوى أدنى بناءً على معيار التفرع. فقط العقد التي تحتوي على مجموعة من الفئات المختلفة يجب أن تكون متفرعة أكثر. أخيراً، توقف خوارزمية شجرة القرار نمو الشجرة بناءً على معيار التقسيم. أبسط معيار تقسيم هو المعيار الذي تنتمي فيه جميع الأمثلة التدريبية إلى نفس الفئة.

3-5 بارامترات تقييم الأداء:

تقييم أداء نماذج التعلم الآلي يتضمن مجموعة من البارامترات التي تتيح فهم كيفية أداء النموذج على مجموعة البيانات المستخدمة.

3-5-1 معدل الخطأ (Error) والدقة (Accuracy):

مقاييس الأداء الأكثر شيوعاً هي معدلات الخطأ والدقة، وذلك في التصنيف الثنائي والتصنيف المتعدد (متعدد الفئات)، حيث أن معدل الخطأ هو نسبة العينات المصنفة بشكل غير صحيح إلى جميع العينات. من ناحية أخرى، فإن الدقة هي نسبة عينات التصنيف الحقيقية إلى جميع العينات، وتُعرف الدقة على أنها عكس معدل الخطأ أي أن: $Accuracy = 1 - Error$.

وتُحسب الدقة بصيغة رياضية كما يلي:

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

حيث:

- TP (True Positives) هي الحالات التي تم التنبؤ بها بشكل صحيح كإيجابية.
- TN (True Negatives) هي الحالات التي تم التنبؤ بها بشكل صحيح كسلبية.
- FP (False Positives) هي الحالات التي تم التنبؤ بها بشكل خاطئ كإيجابية.
- FN (False Negatives) هي الحالات التي تم التنبؤ بها بشكل خاطئ كسلبية.

3-5-2 الضبط والاستدعاء وF1:

عندما يتعين علينا تقييم نموذج ما، فإننا غالباً ما نستخدم معدلات الخطأ والدقة، ولكن ما نركز عليه بشكل أساسي هو مدى موثوقية نموذجنا ومدى مرونته، وكيف يعمل على مجموعة بيانات مختلفة. لا شك أن الدقة معيار مهم للغاية يجب أخذه في الاعتبار، ولكنها لا تقدم دائماً صورة كاملة لأداء النموذج فهي تساهم في فهمنا المتعمق للنموذج وتساعد على تحسينه لكن لتحسين النموذج بشكل فعلي نحتاج إلى النظر في كيفية عمل النموذج على مستوى أعمق. ومع ذلك، لا يتم تحقيق ذلك بمجرد النظر إلى معيار الدقة فقط، على سبيل المثال كيف يدرك البيانات وكيف يمكن التنبؤ بها، وبالتالي للحصول على صورة كاملة للنموذج يجب أن يتم النظر في معايير أخرى مثل الضبط (Precision) والاستدعاء (Recall) وF1 "هي أمثلة على هذه المعايير".

- **الاستدعاء Recall:** معيار هام لأن القيمة العالية له تدل على قيمة منخفضة للنتائج السلبية الخاطئة، وبذلك لن يتم تصنيف رسائل بريد تصيد على أنها رسائل بريد آمنة. فهو يشير إلى قدرة النموذج على التنبؤ بالحالات الإيجابية من بين جميع الحالات الإيجابية الحقيقية:

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

على سبيل المثال، إذا تم تصنيف رسالة آمنة على أنها احتيالية، فهذا يكون له عواقب وخيمة للغاية على النظام.

- **الضبط Precision:** يُظهر جزء الإيجابيات الحقيقية بين العينات التي يتوقع أن تكون موجبة:

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

يعتبر الضبط معيار هام لأن القيمة العالية له تدل على قيمة منخفضة للنتائج الإيجابية الخاطئة، وبذلك لن يفقد المستخدم رسائل البريد الإلكتروني المهمة. على سبيل المثال، عند اكتشاف البريد الإلكتروني العشوائي، تعني النتائج الإيجابية الخاطئة أن البريد الإلكتروني الذي ليس بريداً عشوائياً (منفي حقيقي) تم تحديده على أنه بريد عشوائي (بريد عشوائي متوقع). إذا كانت دقة نموذج الكشف عن الرسائل غير المرغوب فيها عالية، فقد يفقد المستخدم رسائل البريد الإلكتروني المهمة.

- **F1:** درجة F1 هي مقياس يجمع بين Recall و Precision. ببساطة، تجمع F1 بين الضبط والاستدعاء في معيار واحد عن طريق حساب المتوسط التوافقي بين الاثنين:

$$F1 = 2 * \frac{\text{Recall} * \text{Precision}}{\text{Recall} + \text{Precision}}$$

3-5-3 مصفوفة الارتباك (Confusion Matrix):

لحساب معايير التقييم لنموذج التصنيف، نحتاج إلى أربع مجموعات والتي يمكن تمثيلها في مصفوفة الارتباك (Confusion Matrix) كما يوضح الشكل (3.16) وهي: الموجب الحقيقي، الموجب الخاطئ، المنفي الحقيقي والمنفي الخاطئ ويمكن تعريفها كما يلي:

- موجب حقيقي (TP True Positive): عندما كانت القيمة الفعلية للفئة "نعم"، يكون توقع النموذج أيضاً "نعم" (أي توقع صحيح). في دراستنا، عندما تكون القيمة الفعلية للفئة هي يريد آمن ويتم التوقع أنها يريد آمن.
- موجب خاطئ (FP False Positive): عندما كانت القيمة الفعلية للفئة "نعم" لكن النموذج توقع "لا" (أي توقع خاطئ). في دراستنا، عندما تكون القيمة الفعلية للفئة هي يريد آمن ويتم التوقع على أنه يريد تصيد.
- منفي خاطئ (FN False Negative): عندما تكون القيمة الفعلية للفئة "لا"، لكن النموذج توقع "نعم" (أي توقع خاطئ). في دراستنا، عندما تكون القيمة الفعلية هي يريد تصيد ويتم التوقع على أنه يريد آمن.
- منفي حقيقي (TN True Negative): عندما تكون القيمة الفعلية للفئة "لا" وتوقع النموذج "لا" (أي توقع صحيح). في دراستنا، عندما تكون القيمة الفعلية هي يريد تصيد ويتم التوقع أنه يريد تصيد.

		الفئة المتوقعة	
		مثبت	منفي
الفئة الحقيقي	مثبت	موجب حقيقي (TP)	منفي خاطئ (FN)
	منفي	موجب خاطئ (FP)	منفي حقيقي (TN)

الشكل (3-16): مصفوفة الارتباك

3-6 الاستبيان:

يهدف الاستبيان إلى تقييم الآراء والتجارب المختلفة فيما يتعلق بالهجمات المرتبطة برسائل البريد الإلكتروني وفهم مدى الاستعداد لتبني حلول الذكاء الاصطناعي لتعزيز أمن البريد الإلكتروني. تم إنشاء الأسئلة باستخدام استبيان جوجل ووصل عدد الردود إلى أكثر من 178 إجابة، مع الإشارة إلى أن الشريحة المستهدفة في هذا الاستبيان هي بالمعظم من طلاب الجامعات.

1-6-3 مناقشة نتائج الاستبيان:

أظهرت نتائج الاستبيان أن نسبة كبيرة من المشاركين في الاستبيان (حوالي 70%) يدركون أنه ليست كل رسائل البريد الإلكتروني آمنة، وهذا يدل على الوعي لدى نسبة كبيرة من المشاركين حول وجود هجمات مرتبطة برسائل البريد الإلكتروني.

ومن خلال السؤال الثاني والثالث تبين أن 71% من المشاركين اكتشفوا سابقاً رسائل غير مرغوب فيها في بريدهم الوارد، وأن أكثر من نصف المستجيبين وجدوا الرسائل المهمة في صندوق البريد العشوائي الخاص بهم، وهذا يؤكد على ضرورة إيجاد تقنيات أكثر فعالية في تصنيف رسائل البريد الإلكتروني. كذلك تبين في السؤال الرابع أن حوالي نصف المشاركين يقومون بحذف رسائل البريد الإلكتروني العشوائية، ويتجاهلها الثلث، وينقلها الباقون إلى صندوق البريد العشوائي، مما يعكس مدى إدراك المشاركين لخطورة مثل هذه الرسائل.

كما نلاحظ في السؤال الخامس أن حوالي 45% من المستجيبين لا يفتحون روابط من رسائل البريد الإلكتروني المجهولة، وحوالي 30% منهم يفتحون روابط إذا كان مصدر الرسالة معروفاً، و18% منهم لا يفتحون أي رابط بغض النظر عن مصدر الرسالة، وتثبت هذه النسب إلى أن أفراد العينة يدركون خطورة الروابط ويشعرون بالحذر منها.

أظهرت نتائج الاستطلاع في السؤال السادس أن معظم رسائل البريد الإلكتروني غير المرغوبة كانت عبارة عن إعلانات (حوالي 85%) وكان حوالي ثلثي المشاركين على علم بوجود رسائل بريد إلكتروني تصيدية أو تحتوي على روابط لبرامج ضارة، وهذا يعكس معرفة المشاركين بالتصنيفات المختلفة لرسائل البريد الإلكتروني غير المرغوبة.

وفي السؤال السابع لا يعير حوالي ثلثي المشاركين أي اهتمام لرسائل البريد الإلكتروني التي تحتوي على معلومات جاذبة، مما يشير إلى أن نسبة معقولة من المشاركين لا يولون اهتمام لمثل هذه

الرسائل. وفي السؤال الثامن يرى حوالي نصف المشاركين أن البريد الإلكتروني هو وسيلة غير آمنة لإرسال معلومات حساسة. وهذا يؤكد ضرورة وجود طرق تصنيف أكثر فعالية وأمان لتصفية مثل هذه الرسائل قبل وصولها إلى المستخدمين.

وفي السؤال التاسع يعتقد معظم المشاركين (حوالي 90%) أنه يجب أن تكون هناك طرق أكثر فعالية لتصنيف رسائل البريد الإلكتروني، حيث أن حوالي ثلث المشاركين لديهم فكرة عن آلية عمل ودور تطبيقات الذكاء الاصطناعي في مختلف المناحي كما هو موضح في السؤال العاشر.

وأخيراً يظهر السؤال الحادي عشر اعتقاد معظم المشاركين (حوالي 90%) أن تقنيات الذكاء الاصطناعي يمكن أن تلعب دوراً فعالاً في تصنيف رسائل البريد الإلكتروني. وهذا يؤكد على أهمية البحث انطلاقاً من آراء المشاركين في إيجاد خوارزميات الذكاء الاصطناعي المناسبة التي تقوم بتصنيف رسائل البريد الإلكتروني بشكل فعال.

الفصل الرابع

النتائج والمناقشة

4-1 تمهيد:

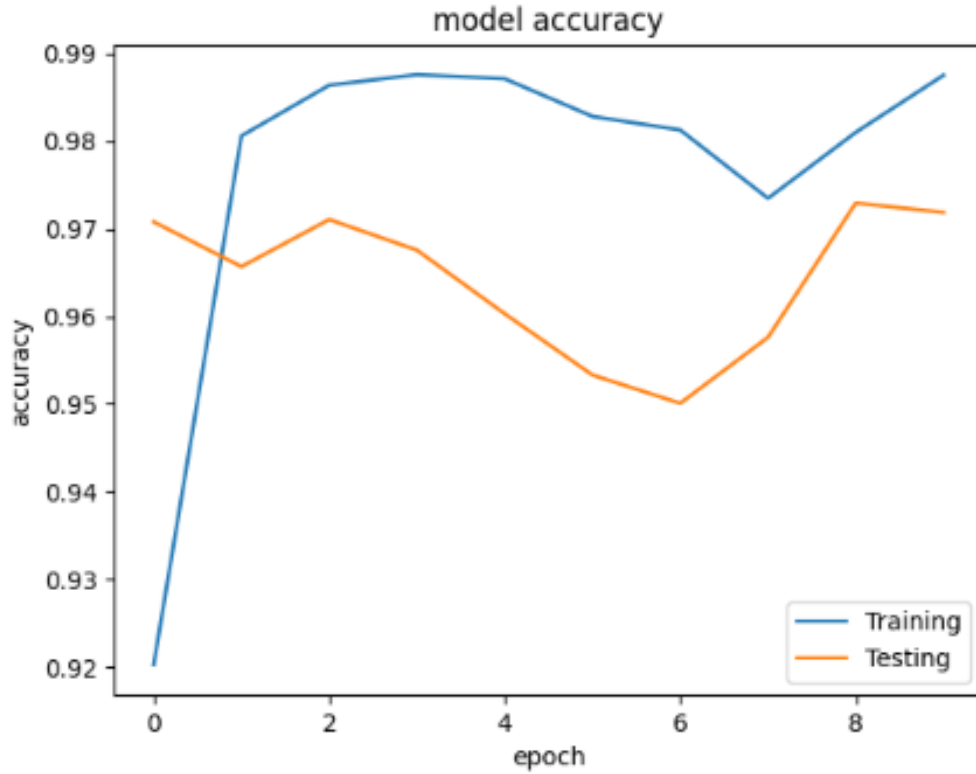
يستعرض هذا الفصل نتائج تقييم النماذج الخاصة بكشف الهجمات على مواقع الويب، مع التركيز على المقاييس الرئيسية مثل الدقة، وتابع الضياع، ومصفوفة الارتباك. وباعتبار أن مجموعة البيانات الخاصة بأشهر الهجمات على مواقع الويب منفصلة عن مجموعة البيانات الخاصة برسائل البريد الإلكتروني بسبب وجود خصائص مختلفة لهذه الهجمات. بالنتيجة، سيكون لدينا تطبيق نماذج مختلفة وتدريبها على كل من مجموعتي البيانات، وسيتم تقييم النتائج الخاصة بكل مجموعة بيانات على حدة، ولكن في النهاية سوف يتم التوصل إلى نتائج مشتركة فيما يخص حماية مواقع الويب من مختلف الهجمات، ليتم في المرحلة الأخيرة كخطوة متممة لهذه الدراسة تطبيق نتائج الدراسة التي تم التوصل إليها على موقع ويب افتراضي. مما يعطي إدارة وتصور كامل حول الحماية المتكاملة لمواقع الويب من الهجمات ويؤكد قابليتها للتطبيق في مواقع الويب الحقيقية.

4-2 تقييم أداء نماذج اكتشاف وتصنيف رسائل البريد الإلكتروني:

النموذج المستخدم لتصنيف رسائل البريد الإلكتروني إلى رسائل بريد إلكتروني سليمة أو تصيدية على مجموعة البيانات (Phishingemails) هو نموذج تعلم عميق للشبكات العصبية التكرارية RNN من نوع الوحدات ذات البوابات المتكررة (GRU) مع الطريقة Bidirectional. عملية تدريب نموذج التعلم العميق تمت وفق 10 epoch، حيث يمثل عدد Epochs (أو "الأدوار") عدد المرات التي يتم فيها تمرير كامل مجموعة البيانات التدريبية عبر شبكة التعلم العميق، وقد تم اختيار قيمة epoch=10 لأنه من خلال هذه القيمة لعدد الأدوار تمكنا من الوصول إلى أعلى دقة سواء بالنسبة لـ Accuracy .Validation or Loss Validation

1-2-4 نتائج الدقة:

بنتيجة عملية التدريب تم الحصول على مخطط لدقة النموذج في حالتي التدريب والاختبار كما يوضح الشكل (1-4).

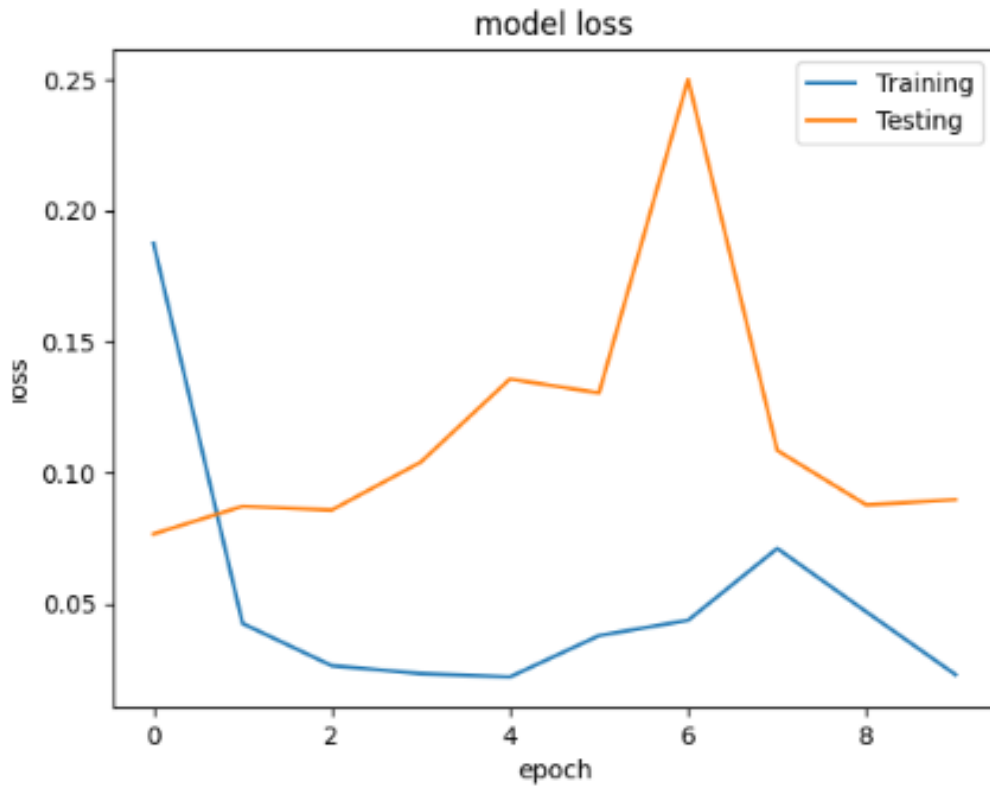


الشكل (1-4): دقة نموذج التعلم العميق بحالتي الاختبار والتدريب

نلاحظ أن قيمة الدقة (دقة اختبار 97% ودقة التدريب 98%)، وهي دقة جداً متقاربة بين Training Accuracy و Validation Accuracy. ولا يوجد مشكلة Overfitting، حيث أن مشكلة Overfitting تحصل عندما تكون دقة التدريب متباعدة عن دقة الاختبار.

2-2-4 تابع الضياع:

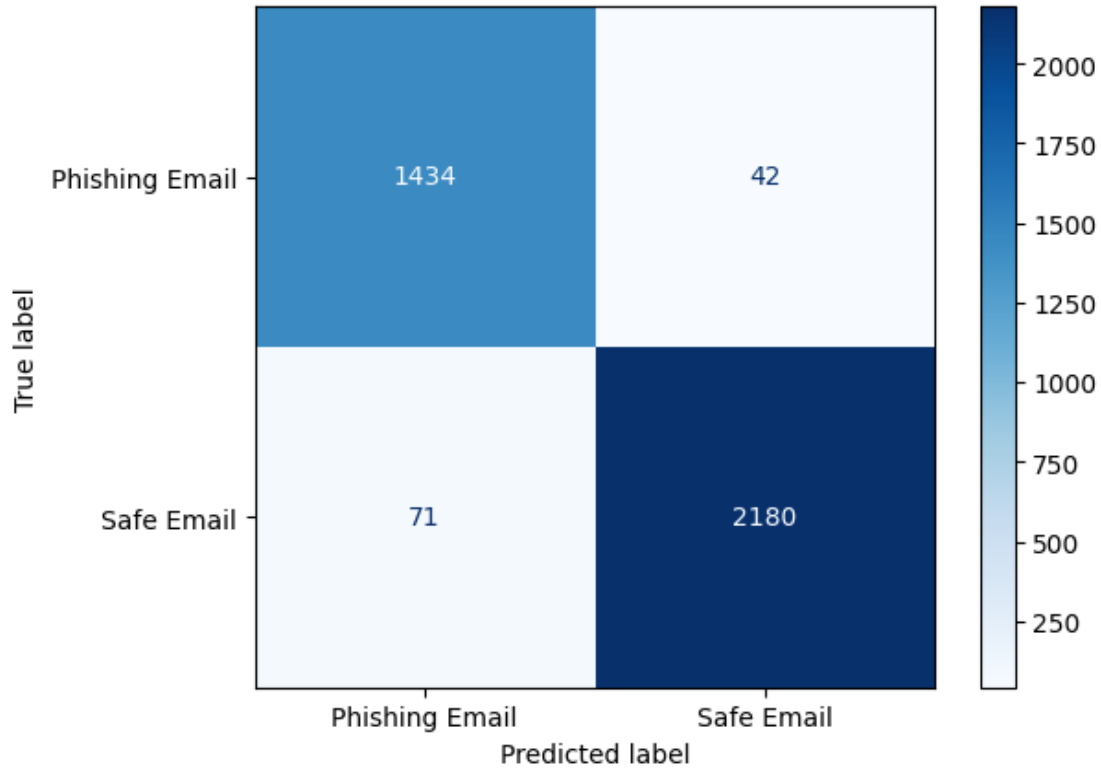
تابع الضياع هو مقياس يستخدم لتقييم أداء نموذج التعلم الآلي من خلال تحديد مقدار الخطأ في التنبؤات التي يقوم بها النموذج. يوضح الشكل (2-4) منحنى الضياع بحالتي التدريب والاختبار. نلاحظ أنه لا يوجد فرق كبير بين قيمتي الضياع بحالتي التدريب والاختبار وإن قيمة الخطأ لا تتجاوز 10% وبالتالي فإن النموذج دقيق ومناسب.



الشكل (2-4): تابع الضياع بحالتي الاختبار والتدريب

3-2-4 مصفوفة الارتباك (Confusion Matrix):

مصفوفة الارتباك التي تم الحصول عليها بنتيجة تقييم أداء النموذج مبينة في الشكل (3-4)، حيث تمثل القيم الموجودة في القطر الرئيسية القيم الصحيحة للتوقعات، بينما تمثل القيم الموجودة في القطر الثانوي التوقعات الخاطئة.



الشكل (3-4): نتيجة مصفوفة الارتباك لتصنيف رسائل البريد

نلاحظ أن عدد الحالات الصحيحة لتصنيف رسائل البريد الإلكتروني الآمنة هي 1434 حالة، وعدد الحالات الصحيحة لتصنيف رسائل البريد الإلكتروني من نوع تصيد هي 2180. بالمقابلة فإن عدد حالات التصنيف الخاطئة لرسائل البريد الآمنة (والتي صُنفت على أنها رسائل تصيد) هو 42 حالة، وعدد حالات التصنيف الخاطئة لرسائل بريد التصيد (والتي صُنفت على أنها رسائل آمنة) هي 71 حالة.

4-2-4 الضبط والاستدعاء وF1:

تظهر النتائج المبينة في الجدول (1-4) قيم الأداء لبارامترات الدقة والضبط والاستدعاء وF1، حيث تتراوح قيم هذه البارامترات بين 96% و98%، وهي نتائج ممتازة وخاصة بالنسبة لقيمة الضبط (Precision) التي تمثل النتائج الصحيحة إيجابياً من بين جميع النتائج التي تم تصنيفها إيجابياً بواسطة النموذج، أي رسائل البريد الآمنة التي تم تصنيفها بشكل صحيح.

الجدول (1-4): نتيجة بارامترات الأداء: الدقة والضبط والاستدعاء وF1

Accuracy	Precision	Recall	F1
0.9718	0.9868	0.9662	0.9764

3-4 تقييم أداء نماذج اكتشاف وتصنيف الهجمات على مواقع الويب:

سوف يتم تطبيق مجموعة من خوارزميات التعلم الآلي والتعلم العميق على مجموعة البيانات (ids-intrusion) التي تحتوي على أبرز الهجمات على مواقع الويب. من خوارزميات التعلم الآلي التي يتم تطبيقها، خوارزمية شجرة القرار (Decision Tree) وخوارزمية الغابة العشوائية (Random Forest) وخوارزمية (XGBClassifier)، أما من خوارزميات التعلم العميق فيتم تطبيق المصنف (MLP).

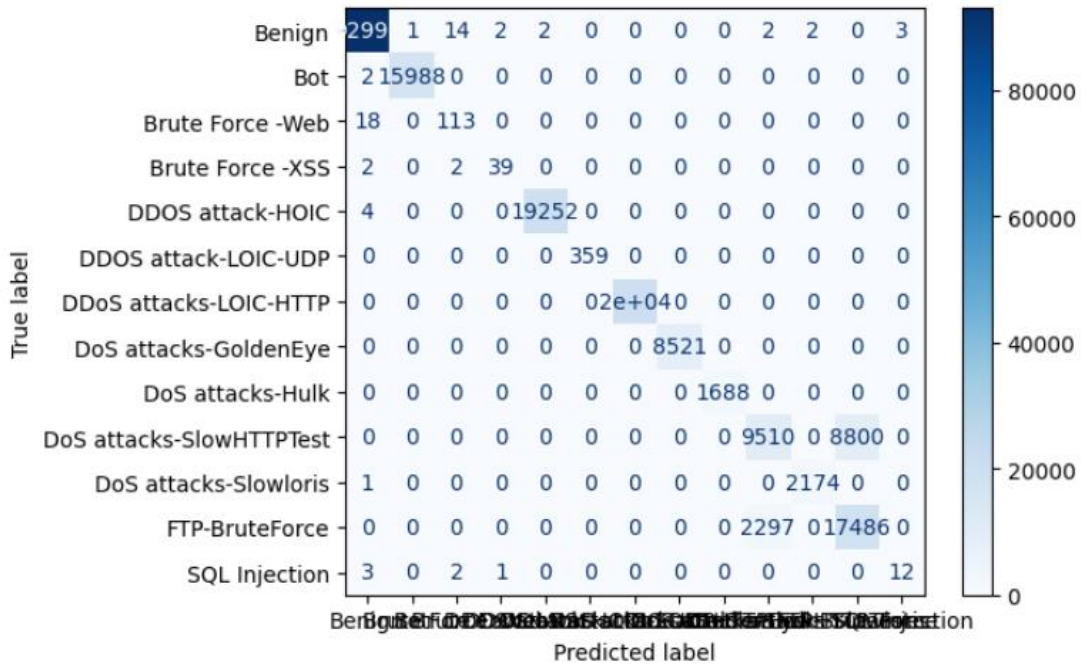
بنتيجة تدريب الخوارزميات المذكورة على مجموعة البيانات المستخدمة نحصل على نتائج تقييم الأداء المبينة في الجدول (2-4). حيث يتبين أن أداء خوارزمية الغابة العشوائية هي أفضل خوارزمية من بين الخوارزميات المطبقة حيث لديها أعلى قيم لبارامترات الأداء المدروسة. وبشكل مماثل تقريباً لخوارزمية الغابة العشوائية تكون نتائج خوارزمية شجرة القرار وخوارزمية XGBClassifier جيدة. ولكن نتائج خوارزميات مصنف (MLP) تعتبر سيئة حيث لا تتجاوز الدقة فيها 65%.

الجدول (2-4): نتائج تقييم الأداء لمختلف الخوارزميات

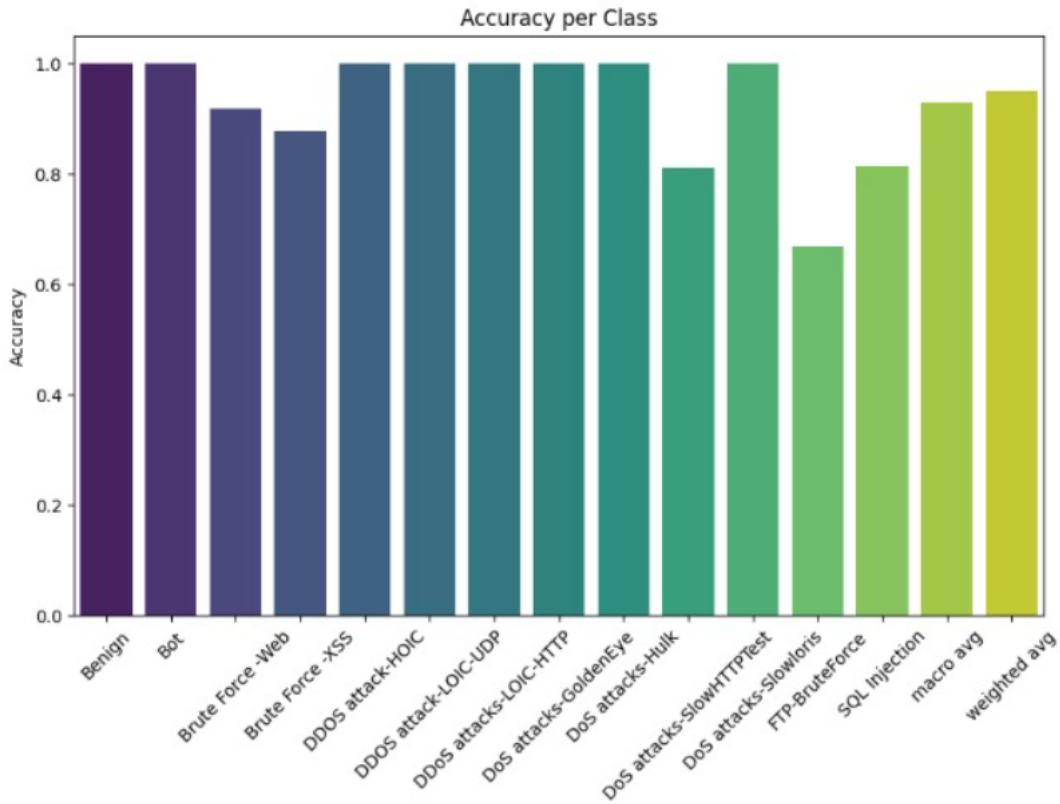
	Accuracy	Precision	Recall	F1
Random Forest	0.94474	0.956432	0.94474	0.946863
Decision Tree	0.94452	0.9563846	0.94452	0.94659
XGBClassifier	0.94391	0.95549	0.9439127	0.94598
MLPClassifier	0.64736	0.99488	0.64736	0.771815

بالنتيجة فإن النتائج التي يمكن اعتمادها لكشف الهجمات على مواقع الويب هي لخوارزمية الغابة العشوائية.

يبين الشكل (4-4) مصفوفة الارتباك لخوارزمية الغابة العشوائية، حيث يمثل القطر الرئيسي القيم الصحيحة لتصنيفات الهجمات بينما تمثل بقية قيم الأسطر التصنيفات الخاطئة لكل من حركة المرور الطبيعية وكل أصناف الهجمات المدروسة.



الشكل (4-4): نتيجة مصفوفة الارتباك لكشف الهجمات



الشكل (5-4): قيمة الدقة لكل صنف من أصناف مجموعة البيانات

يوضح الشكل (4-5) قيم الدقة لكل صنف من أصناف مجموعة البيانات بالنسبة لخوارزمية الغابة العشوائية، حيث يتبين أن قيمة الدقة في تصنيف حركة المرور الطبيعية (Benign) هي مئة بالمئة بينما تتراوح قيمتها بين 65% و100% بالنسبة لأصناف الهجمات المختلفة.

4-4 الدراسة التطبيقية لكشف وحماية موقع ويب من الهجمات الخبيثة:

في هذا الجزء يتم الانتقال إلى الدراسة التطبيقية لحماية لكشف الهجمات على موقع ويب بشكل فعلي وتأمين الحماية المطلوبة بشكل حقيقي، وذلك بعد أن تم الانتهاء من الدراسة التحليلية واستنتاج أفضل خوارزميات الذكاء الاصطناعي التي تقوم بدور فعال في كشف وتصنيف الهجمات.

بالتفصيل، نقوم بتحميل موقع ويب معين موجود على شبكة الانترنت وتحملته على جهاز الحاسب لنقوم بعمليات الاختبار بشكل محلي على الموقع المحدد، يتم ربط موقع الويب مع خوارزميات الذكاء الاصطناعي المدروسة سابقاً (خوارزمية RNN مع GRU و Bidirectional) لكشف الهجمات وتصنيف هجمات رسائل البريد الإلكتروني، إضافة إلى خوارزمية الغابة العشوائية لكشف الهجمات الشهيرة على موقع الويب الافتراضي).

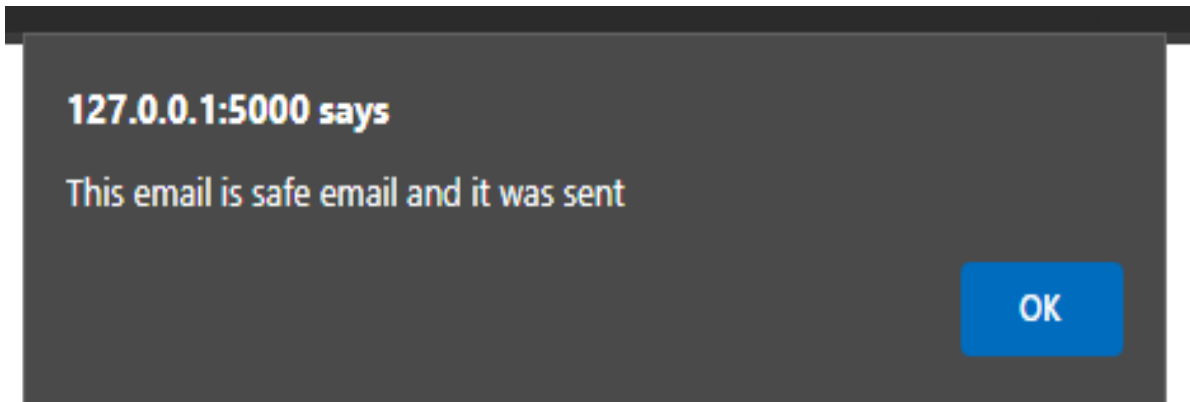
4-4-1 اختبار كشف وتصنيف رسائل البريد الإلكتروني:

بالنسبة لتصنيف رسائل البريد الإلكتروني يوجد في الموقع قسم للاتصال (Contacts)، حيث يمكن لأي شخص إرسال بريد إلكتروني، قد يكون هذا البريد الإلكتروني عبارة عن بريد إلكتروني للتصيد الاحتمالي، لذا يتم استخدام نموذج التعلم العميق للكشف عن رسائل البريد الإلكتروني وتصنيفها فيما إذا كانت تصيد أو آمنة.

بدايةً، عندما يقوم المستخدم بإدخال رسالة البريد الإلكتروني سليمة عند الضغط على Submit وإرسال الرسالة كما في الشكل (4-6)، يتم فحص هذه الرسالة من قبل نموذج التعلم الآلي العميق من أجل تصنيف وتحديد نوع هذه الرسالة، نتيجة تصنيف الرسالة التي سوف نحصل عليها من قبل النموذج بأنها رسالة سليمة كما هو مبين في الشكل (4-7).

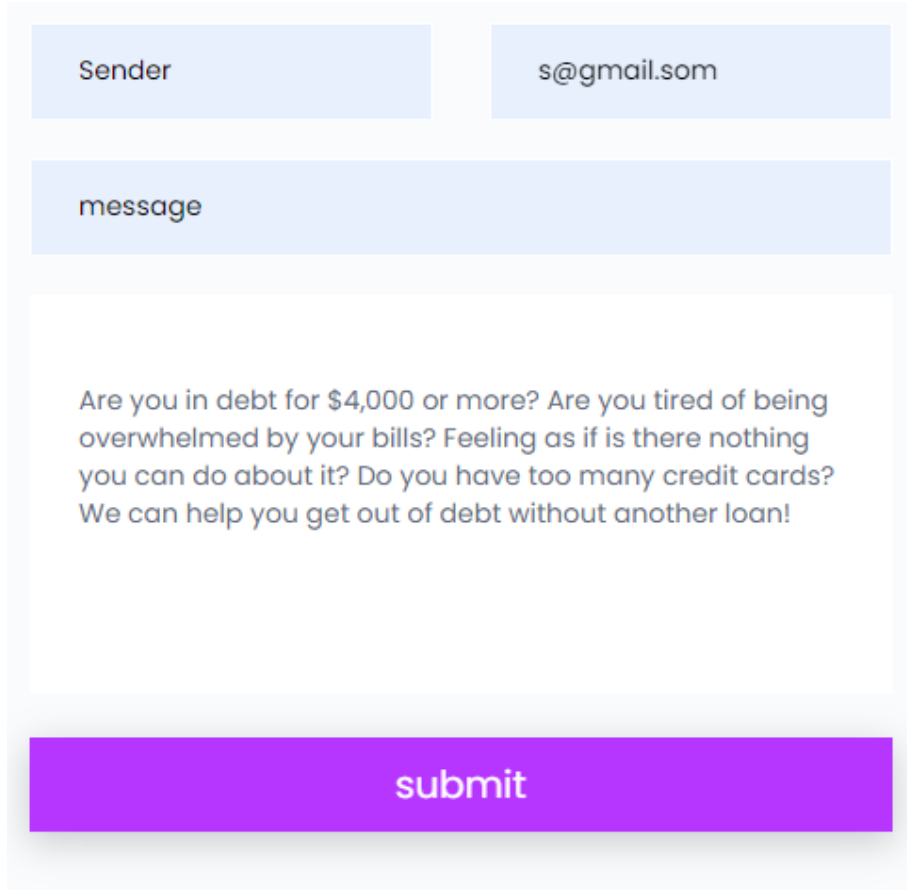
Sender	s@gmail.com
message	
<p>summer at enron hi vince : if you or your human resources department tries to reach me, please call me at new home phone (713) 647 - 7161 or email me at ekao @ uh. edu , or eckao @ aol . com monday , 5 / 15 will be my moving day. otherwise, I can be reached at my home phone. regards, ed</p>	
<input type="submit" value="submit"/>	

الشكل (4-6): اختبار إرسال رسالة بريد إلكتروني سليمة



الشكل (4-7): نتيجة تصنيف رسالة البريد الإلكتروني السليمة من قبل نموذج الذكاء

بالمقابل، عندما يقوم المستخدم بإدخال رسالة البريد الإلكتروني من نوع تصيد كما في الشكل (8-4).



Sender s@gmail.com

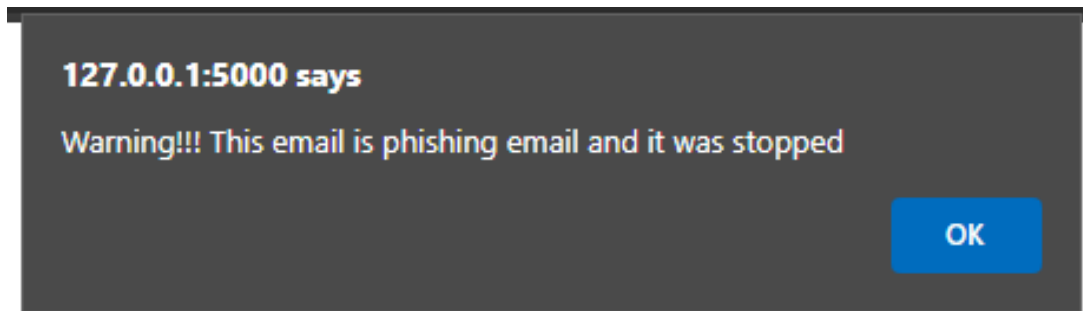
message

Are you in debt for \$4,000 or more? Are you tired of being overwhelmed by your bills? Feeling as if is there nothing you can do about it? Do you have too many credit cards? We can help you get out of debt without another loan!

submit

الشكل (8-4): اختبار إرسال رسالة بريد إلكتروني من نوع تصيد

النتيجة التي يتم الحصول عليها هي كثف نموذج الذكاء لرسالة التصيد كما يوضح الشكل (9-4).

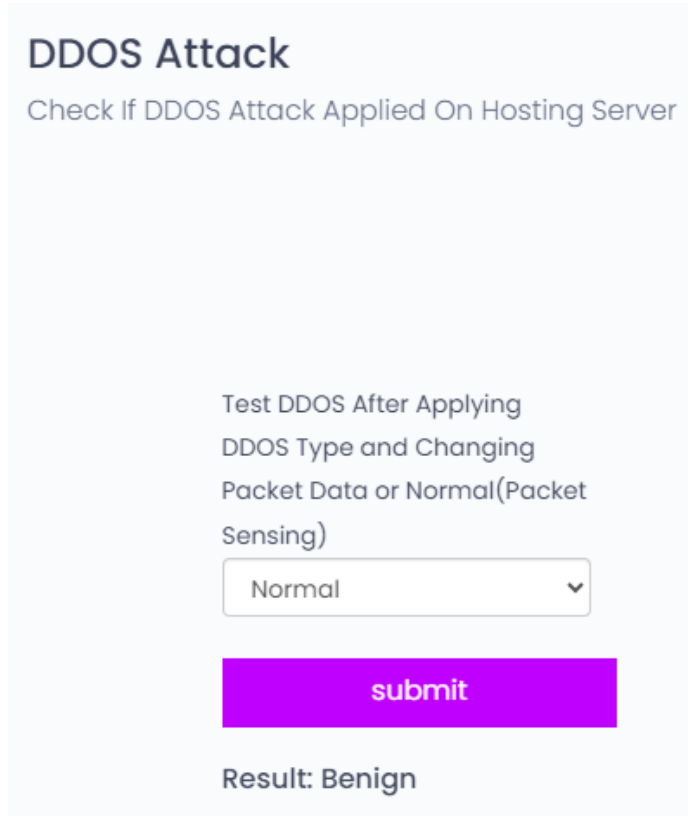


الشكل (9-4): نتيجة تصنيف رسالة البريد الإلكتروني من نوع تصيد من قبل نموذج الذكاء

بالنتيجة، يقوم نموذج التعلم العميق المستخدم بتصنيف رسائل البريد الإلكتروني (السليمة والتصيد) بشكل صحيح، وهذه نتيجة طبيعية كون دقة تصنيف هذا النموذج بنتيجة الدراسة التحليلية هي حوالي 97%، أي أنها من أجل 100 رسالة بريد إلكتروني، فإنه يقوم بتصنيف 97 رسالة بشكل صحيح و3 رسائل بشكل خاطئ.

2-4-4 اختبار كشف وتصنيف الهجمات:

بشكل مماثل لطريقة اختبار تصنيف رسائل البريد الإلكتروني، يوجد في قسم الاتصال (Contacts) ضمن موقع الويب أداة لاكتشاف ما إذا كان هناك هجوم على خادم الاستضافة الخاص بموقع الويب أم لا. لاختبار خوارزمية التعلم الآلي المستخدمة لكشف الهجمات على موقع الويب المحدد، يمكن للمستخدم استخدام الوضع العادي (أي حركة المرور الطبيعية)، وسوف تستشعر الأداة الحزم المستلمة للكشف عن نوع الهجوم ويكتشف نموذج الغابة العشوائية أنه لا يوجد هجوم (Benign) في حركة المرور كما هو موضح في الشكل (4-10).



DDOS Attack
Check If DDOS Attack Applied On Hosting Server

Test DDOS After Applying
DDOS Type and Changing
Packet Data or Normal(Packet
Sensing)

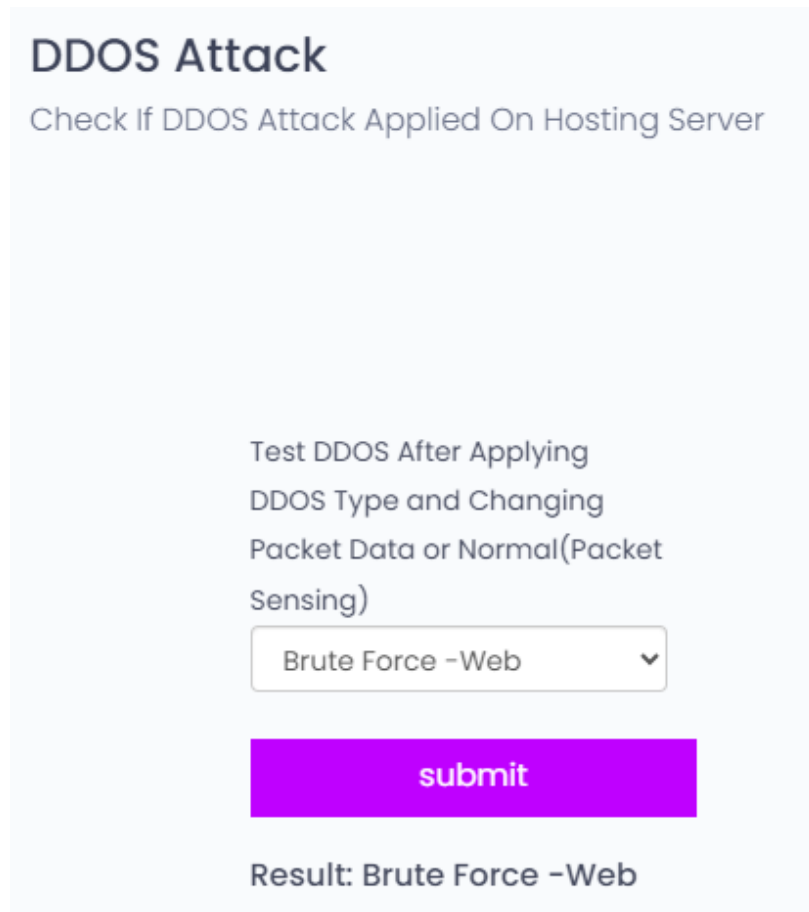
Normal

submit

Result: Benign

الشكل (4-10): اختبار توليد وكشف حركة المرور الطبيعية

بالمقابل، يمكن للمستخدم تحديد أي نوع من الهجوم وستقوم الأداة بتغيير بيانات الحزمة وإدخال الحزمة إلى النموذج لاكتشافها. تمت إضافة هذه الخطوة لتوضيح كيفية اكتشاف الخادم لأنواع الهجوم، على سبيل المثال عند توليد الأداة لهجوم من نوع Brute force بالضغط على Submit ليقوم بعدها نموذج تعلم الآلي المستخدم بكشف أن هذا الهجوم هو من نوع Brute force كما هو موضح في الشكل (11-4). وبنفس الطريقة من أجل توليد هجوم من نوع آخر هو (SQL Injection) كما هو موضح في الشكل (12-4).



DDOS Attack
Check If DDOS Attack Applied On Hosting Server

Test DDOS After Applying
DDOS Type and Changing
Packet Data or Normal(Packet
Sensing)

Brute Force -Web

submit

Result: Brute Force -Web

الشكل (11-4): اختبار توليد وكشف هجوم من Brute force

DDOS Attack

Check If DDOS Attack Applied On Hosting Server

Test DDOS After Applying
DDOS Type and Changing
Packet Data or Normal(Packet
Sensing)

SQL Injection

submit

Result: SQL Injection

الشكل (4-12): اختبار توليد وكشف هجوم من SQL Injection

الفصل الخامس

الاستنتاجات والتوصيات والمقترحات

1-5 الاستنتاجات:

في ختام هذا البحث، تم استكشاف أداء نماذج الذكاء الاصطناعي في اكتشاف وتصنيف الهجمات على مواقع الويب ورسائل البريد الإلكتروني. أظهرت النتائج قدرة ملحوظة على التعرف على الأنماط الخبيثة بدقة عالية تتجاوز 95% في معظم الحالات، مما يعكس فعالية النماذج المستخدمة في هذا السياق. ومن الملفت للانتباه عدم وجود مشاكل فرط التدريب (Overfitting)، مما يعزز موثوقية النتائج ويؤكد جودة النماذج المدروسة.

تم تطبيق هذه النماذج عملياً على موقع ويب افتراضي، وأظهرت نتائج إيجابية في اكتشاف الهجمات بدقة مشابهة لتلك التي حصلت خلال التقييمات الاختبارية. هذا يشير إلى إمكانية استخدام هذه التقنيات في بيئات الإنترنت الحقيقية لتعزيز الأمن السيبراني وحماية المواقع من التهديدات الخبيثة.

بالإضافة إلى ذلك، يسלט البحث الضوء على أهمية تطوير وتبني التقنيات الحديثة مثل الذكاء الاصطناعي في اكتشاف وتصنيف الهجمات الخبيثة على موقع الويب. وتشير النتائج الإيجابية المحققة في هذا العمل إلى أن هذه التقنيات قد تكون ذات قيمة كبيرة في مواجهة التهديدات السيبرانية المتزايدة في العالم الرقمي الحديث. بالنهاية يمكن تلخيص أبرز ما تم التوصل إليه في هذه الأطروحة فيما يلي:

- إيجاد نموذج تعلم آلي (مصنف الغابة العشوائية) للكشف عن الهجمات الشائعة على مواقع الويب، يتميز هذا النموذج بدقة إجمالية حوالي 94%، وبدقة تصل إلى 100% في تصنيف حركة المرور الطبيعية.

- بناء نموذج تعلم عميق (شبكات عصبية تكرارية RNN من نوع GRU مع طريقة Bidirectional) لتصنيف رسائل البريد الإلكتروني، يتميز هذا النموذج بدقة تصل إلى حوالي 97% مع عدم وجود مشكلة الضبط الزائد (Overfitting).

- تطبيق النماذج التي تم التوصل إليها في الدراسة التحليلية على موقع ويب افتراضي، من خلال ربط نماذج الذكاء الصناعي المدروسة مع مواقع الويب، واستخدام هذه النماذج لكشف الهجمات وتصنيف حركة المرور الطبيعية. وقد أثبتت نتائج التطبيق فعالية النماذج المدروسة في عملية الكشف والتصنيف لأنماط الهجوم المختلفة.

5-2 التوصيات المستقبلية:

بناءً على نتائج هذا البحث وتجاربه، نقدّم بعض التوصيات المستقبلية التي يمكن اعتبارها للأبحاث المقبلة في هذا المجال:

- توسيع مجموعات البيانات: ينبغي توسيع مجموعات البيانات المستخدمة في التدريب والاختبار لتضمن تنوعاً أكبر من السيناريوهات الخبيثة والسليمة، مما يزيد من قدرة النماذج على التعرف على الهجمات المستقبلية.
- تحسين أداء نماذج الذكاء الاصطناعي: استكشاف تقنيات التعلم العميق والتعلم الآلي، وتحسين دقتها وفعاليتها في اكتشاف الهجمات بدقة عالية.
- تكامل الأنظمة الأمنية: يمكن دمج نماذج الذكاء الاصطناعي في أنظمة الأمان الشاملة لمواقع الويب لتعزيز الدفاع عنها ضد التهديدات السيبرانية.

3-5 المقترحات:

ننصح بالاعتماد على نتائج هذه الدراسة فيما يتعلق بالكشف عن أبرز الهجمات السيبرانية على مواقع الويب إضافة إلى رسائل البريد الإلكتروني التصيدية كونها أثبتت فعاليتها بنتيجة الدراستين التحليلية والتطبيقية، ونوصي للأعمال المستقبلية بتطبيق طرق عملية لمنع هذه الهجمات عند اكتشافها من قبل النماذج المدروسة، من المحتمل أن تقوم طرق المنع على إسقاط الرزم من نوع هجوم ومنعها من الوصول إلى موقع الويب، أو يمكن اتخاذ أي إجراءات مناسبة لتحقيق الحماية اللازمة عند اكتشاف الهجمات.

المراجع:

- [1] A. J. Hacker and I. CISSP, "Importance of web application firewall technology for protecting web-based resources," ICSA Labs an Independent Verizon Business, 2008.
- [2] Anthi E, Williams L, Rhode M, Burnap P, Wedgbury A. Adversarial attacks on machine learning cybersecurity defences in industrial control systems. *Journal of Information Security and Applications*. 2021 May 1;58:102717.
- [3] S. Saleem, M. Sheeraz, M. Hanif and U. Farooq, "Web Server Attack Detection using Machine Learning," 2020 International Conference on Cyber Warfare and Security (ICCWS), Islamabad, Pakistan, 2020, pp. 1-7, doi: 10.1109/ICCWS48432.2020.9292393.
- [4] Sharma, Sushant & Zavarisky, Pavol & Butakov, Sergey. (2020). Machine Learning based Intrusion Detection System for Web-Based Attacks. 227-230. 10.1109/BigDataSecurity-HPSC-IDS49724.2020.00048.
- [5] Vadhil, Fatimetou & Salihi, Mohamed & Nanne, Mohamedade. (2024). Machine learning-based intrusion detection system for detecting web attacks. *IAES International Journal of Artificial Intelligence (IJ-AI)*. 13. 711. 10.11591/ijai.v13.i1.pp711-721.
- [6] Baklizi, Mahmoud & Atoum, Issa & Alkhazaleh, Mohammad & Kanaker, Hasan & Abdullah, Nibras & Al-Wesabi, Ola & Otoom, Ahmed. (2024). Web Attack Intrusion Detection System Using Machine Learning Techniques. *International Journal of Online and Biomedical Engineering (iJOE)*. 20. 24-38. 10.3991/ijoe.v20i03.45249.

- [7] Choras, Michal & Kozik, Rafal. (2015). Machine learning techniques applied to detect cyber attacks on web applications. *Logic Journal of IGPL*. 23. 45-56. 10.1093/jigpal/jzu038.
- [8] Hoang, Dau & Trang, Ninh & Hung, Nguyễn. (2022). A Survey of Tools and Techniques for Web Attack Detection. *Special Issue CS (15) 2022*. 109-118. 10.54654/isj.v1i15.852.
- [9] Calzarossa, Maria Carla & Giudici, Paolo & Zieni, Rasha. (2023). Explainable machine learning for phishing feature detection. *Quality and Reliability Engineering International*. 40. 10.1002/qre.3411.
- [10] Odiaga, Gloria. (2023). Assessment of Existing Cyber-Attack Detection Models for Web-Based Systems. 15. 070–089. 10.30574/gjeta.2023.15.1.0075.
- [11] Dawadi, Babu R., Bibek Adhikari, and Devesh Kumar Srivastava. 2023. "Deep Learning Technique-Enabled Web Application Firewall for the Detection of Web Attacks" *Sensors* 23, no. 4: 2073 .
- [12] Kurnaz, Sefer & Gwad, Wisam. (2018). Deep Auto-Encoder Neural Network for Phishing Website Classification. 68-72.
- [13] Ghiasi M, Niknam T, Wang Z, Mehrandezh M, Dehghani M, Ghadimi N. A comprehensive review of cyber-attacks and defense mechanisms for improving security in smart grid energy systems: Past, present and future. *Electric Power Systems Research*. 2023 Feb 1;215:108975 .
- [14] Bhamare D, Zolanvari M, Erbad A, Jain R, Khan K, Meskin N. Cybersecurity for industrial control systems: A survey. *computers & security*. 2020 Feb 1;89:101677 .
- [15] Zhao N, Zhao X, Chen M, Zong G, Zhang H. Resilient Distributed Event-Triggered Platooning Control of Connected Vehicles Under Denial-of-

- Service Attacks. *IEEE Transactions on Intelligent Transportation Systems*. 2023 Mar 6.
- [16] Hall RC, Hoppa MA, Hu YH. An Empirical Study of Password Policy Compliance. In *Journal of the Colloquium for Information Systems Security Education* 2023 Mar 8 (Vol. 10, No. 1, pp. 8-8).
- [17] Alarfaj FK, Khan NA. Enhancing the Performance of SQL Injection Attack Detection through Probabilistic Neural Networks. *Applied Sciences*. 2023 Mar 29;13(7):4365.
- [18] Hu T, Xu C, Zhang S, Tao S, Li L. Cross-site scripting detection with two-channel feature fusion embedded in self-attention mechanism. *Computers & Security*. 2023 Jan 1;124:102990.
- [19] Yadav MK, Khan M. Introduction to Web Terminology and Web Application Attacks. *Journal of Web Development and Web Designing*. 2023 Jan 19;8(1):1-2.
- [20] Wang Z, Li Y, Wu S, Zhou Y, Yang L, Xu Y, Zhang T, Pan Q. A survey on cybersecurity attacks and defenses for unmanned aerial systems. *Journal of Systems Architecture*. 2023 Apr 3:102870.
- [21] Tsompanoglou P, Iliadis A, Kantelis K, Petridou S, Nicopolitidis P. Countermeasuring MITM Attacks in Solar-Powered PON-Based FiWi Access Networks. *Electronics*. 2023 Feb 20;12(4):1052.
- [22] Tufan A, Tuna G. Benefits of Information Security Awareness Training Against Phishing Attacks: A Field Study. In *Handbook of Research on Cybersecurity Risk in Contemporary Business Systems* 2023 (pp. 49-78). IGI Global.
- [23] Manning, R., & Aaron, G. (2015). Phishing Activity Trends Report. Anti-Phishing Work Group, Tech. Rep. 1st -3rd Quarter.

- [24] Al-Momani, A., Gupta, B. B., Atawneh, S., Meulenberg, A., & Al-Momani, E. (2013). A survey of phishing email filtering techniques. *Communications Surveys & Tutorials, IEEE*, 15 (4), 2070-2090.
- [25] Cao, Y., Han, W., & Le, Y. (2008). Anti-phishing based on automated individual white-list. In *Proceedings of the 4th ACM workshop on Digital identity management* (pp. 51-60).
- [26] Adida, B., Chau, D., Hohenberger, S., & Rivest, R. L. (2006). Lightweight email signatures. In *Security and Cryptography for Networks* (pp. 288-302). Springer Berlin Heidelberg.
- [27] Ramanathan, V., & Wechsler, H. (2012). phishGILLNET—phishing detection methodology using probabilistic latent semantic analysis, AdaBoost, and co-training. *EURASIP Journal on Information Security*, 1-22.
- [28] Alpaydin, E. (2004). *Introduction To Machine Learning*. S.L.: Mit Press.
- Andrew Carlson, Justin Betteridge, and Bryan Kisiel. (2010). Toward an Architecture for Never-Ending Language Learning. In *AAAI*, pages 1306-1313.
- [29] Oliveira N, Praça I, Maia E, Sousa O. Intelligent cyber-attack detection and classification for network-based intrusion detection systems. *Applied Sciences*. 2021 Feb 13;11(4):1674.

روابط تحميل:

- [30] <https://www.kaggle.com/datasets/subhajournal/phishingemails>
- [31] <https://www.kaggle.com/datasets/solarmainframe/ids-intrusion-csv>

الملاحق:

ملحق (أ): رابط الاستبيان

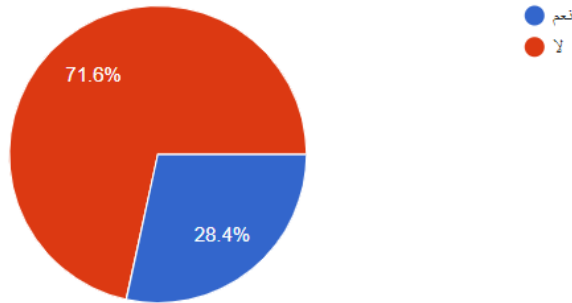
<https://docs.google.com/forms/d/1QOsDMfGTN3cKhGW3GZa5lXXT3cx1BES1w4QlAdTtu8I/edit#responses>

ملحق (ب): نتائج الاستبيان

1- هل تتعامل مع جميع رسائل البريد الإلكتروني الواردة على أنها سليمة وآمنة؟

هل تتعامل مع جميع رسائل البريد الإلكتروني الواردة على أنها سليمة وآمنة؟

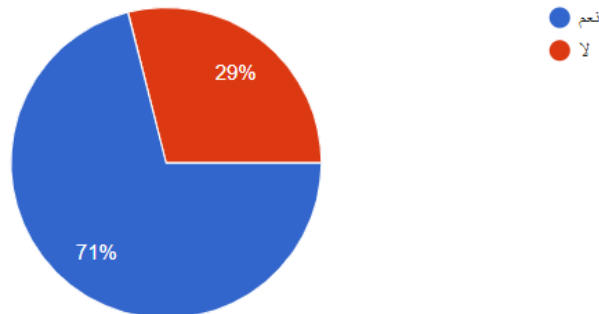
183 ردًا



2- هل سبق ووجدت رسالة بريد إلكتروني غير مرغوبة (Spam) في صندوق الوارد؟

هل سبق ووجدت رسالة بريد إلكتروني غير مرغوبة (Spam) في صندوق الوارد؟

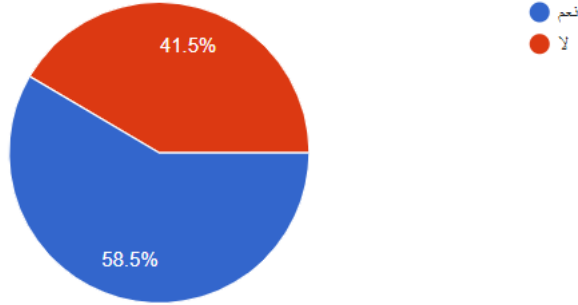
183 ردًا



3- هل سبق ووجدت رسالة هامة في صندوق Spam؟

هل سبق ووجدت رسالة هامة في صندوق Spam؟

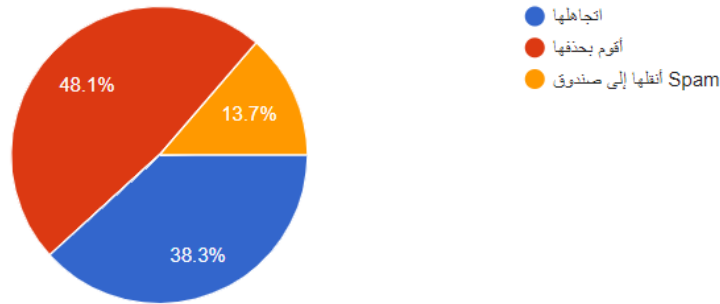
183 رأياً



4- كيف تتعامل مع رسائل البريد الإلكتروني الغير مرغوبة؟

كيف تتعامل مع رسائل البريد الإلكتروني الغير مرغوبة؟

183 رأياً



5- ماذا تفعل عندما تجد رسالة بريد إلكتروني تحتوي على روابط (Links)؟

ماذا تفعل عندما تجد رسالة بريد إلكتروني تحتوي على روابط (Links)؟

183 رأياً

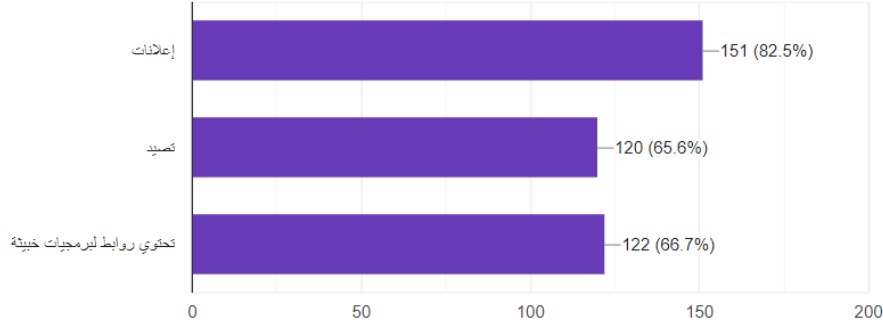


6- ما هي أنواع رسائل البريد الإلكتروني غير المرغوبة التي تعرفها؟



ما هي أنواع رسائل البريد الإلكتروني الغير مرغوبة التي تعرفها؟

183 ردًا

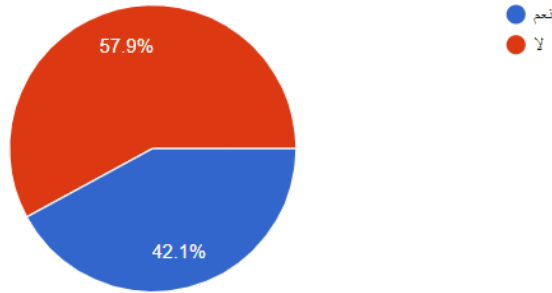


7- هل من الممكن أن تعبر اهتماماً أو تصدق رسائل البريد الإلكتروني التي تحتوي على معلومات

جاذبة؟

هل من الممكن أن تعبر اهتماماً أو تصدق رسائل البريد الإلكتروني التي تحتوي على معلومات جاذبة؟

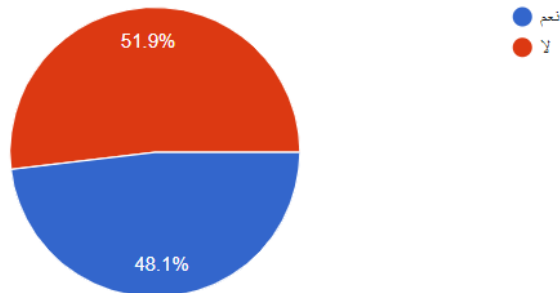
183 ردًا



8- هل ترى أن البريد الإلكتروني وسيلة آمنة لإرسال رسائل هامة وتحتوي معلومات حساسة؟

هل ترى أن البريد الإلكتروني وسيلة آمنة لإرسال رسائل هامة وتحتوي معلومات حساسة؟

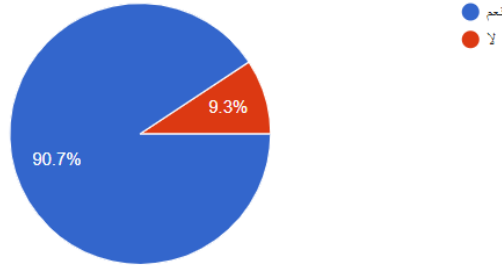
183 ردًا



9- هل تعتقد أنه من الضروري وجود طرق أكثر فعالية لتصنيف رسائل البريد الإلكتروني؟

هل تعتقد أنه من الضروري وجود طرق أكثر فعالية لتصنيف رسائل البريد الإلكتروني؟

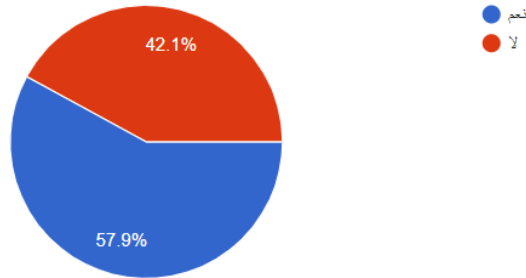
183 رأياً



10- هل لديك فكرة أو معلومات عن آلية عمل ودور تطبيقات الذكاء الاصطناعي؟

هل لديك فكرة أو معلومات عن آلية عمل ودور تطبيقات الذكاء الاصطناعي؟

183 رأياً



11- هل تعتقد أن تقنيات الذكاء الاصطناعي يمكن أن تقوم بدور فاعل في تصنيف وكشف الإيميلات غير المرغوبة؟

هل تعتقد أن تقنيات الذكاء الاصطناعي يمكن أن تقوم بدور فاعل في تصنيف وكشف الإيميلات غير المرغوبة؟

183 رأياً

