

Analyse De Phishing Et Contres Mesures

FOUCHÉ Stanislas stanislasfouche@gmail.com

PRIOU Antoine priou2002@gmail.com

Tuteur : KHOUKHI Lyes

Abstract

Cet article a pour objectif de tester l'efficacité actuelle des modèles de langage de grande taille (LLMs) pour la détection de phishing par mail. Notre hypothèse est que les LLMs pourraient remplacer les méthodes traditionnelles dans ce domaine. Pour tester cette hypothèse, nous avons développé une application permettant aux utilisateurs de transférer un e-mail suspect à une adresse dédiée, où un LLM analyse automatiquement son contenu. En retour, le système fournit un score de probabilité indiquant si l'e-mail est considéré comme un phishing ou non. Une évaluation expérimentale est menée sur un jeu de données afin d'analyser les performances du modèle en termes de précision, de rappel et de robustesse face aux tentatives d'évasion. Enfin, nous discutons des limites de l'approche, notamment la sensibilité aux faux positifs et la nécessité d'adaptation continue face aux nouvelles stratégies d'attaque.

1 Introduction

En 2023, 16.73 millions d'incidents de cybersécurité liés à des attaques de phishing ont été recensés, causant des pertes estimées à plusieurs milliards de dollars selon Statista Research Department [26]. Ces attaques exploitent des failles humaines pour voler des informations sensibles telles que des identifiants, des données bancaires ou encore des documents confidentiels.

Le phishing est aujourd'hui une menace globale, touchant aussi bien des personnalités importantes que des réseaux entiers comme les entreprises avec une approche moins précise mais plus large. Ses conséquences peuvent être dévastatrices : pertes financières, atteinte à la vie privée, ou encore déstabilisation d'organisations entières. Face à ce danger croissant, il est crucial de comprendre son fonctionnement, ses impacts et, surtout, les moyens de s'en protéger.

Au fil des années, différentes méthodes de détection ont pu voir le jour, allant de méthodes traditionnelles jusqu'à l'utilisation de solutions basées sur l'intelligence artificielle. Avec l'arrivée des modèles de langage de grande taille (LLMs), une question se pose : ces modèles peuvent-ils remplacer les méthodes

traditionnelles dans le cadre de la détection de phishing ? Cet article a pour objectif d'analyser les performances actuelles d'un LLM pour la détection de phishing en les comparant aux méthodes existantes.

2 State of Art

2.1 Structure des attaques

L'approche traditionnelle du phishing consiste à envoyer un email, un SMS, ou à passer un appel pour inciter la victime à fournir volontairement des informations sensibles. Cependant, cette méthode a considérablement évolué avec le temps. Elle repose désormais sur le principe de l'ingénierie sociale, qui consiste à étudier le profil de la cible afin de la manipuler psychologiquement.

Les cybercriminels exploitent ainsi diverses techniques pour gagner la confiance de leurs victimes avec des méthodes différentes selon les cibles visées. On peut distinguer plusieurs types de phishing :

- phishing par e-mail : technique classique la plus courante, usurpant des entités légitimes pour inciter la victime à divulguer des informations sensibles via une redirection vers de faux établissements.
- spear phishing: utilise des informations personnalisées pour rendre l'attaque plus crédible
- whaling: vise des personnalités importantes ou des entreprises
- vishing : hameçonnage par téléphone
- Pharming : redirection du DNS vers un site frauduleux
- Clone phishing : copie de mail légitime avec pièces-jointes malveillantes

Cet article se concentre spécifiquement sur l'analyse et la détection du phishing par email. Bien que le phishing partage des similarités avec les spams, notamment en raison de leur diffusion massive, il s'en distingue par sa finalité malveillante et son ciblage souvent très précis. Contrairement aux spams, dont l'objectif est généralement publicitaire ou promotionnel, le phishing vise à tromper la victime afin de lui soutirer des informations sensibles, telles que des identifiants, des

informations bancaires ou des données personnelles. Si tous les types de phishing peuvent effectivement être classifiés comme des formes de spam, il est crucial de noter que l'inverse n'est pas vrai : tous les spams ne relèvent pas du phishing, car leur intention n'est pas toujours de nuire directement à la victime.

2.2 Méthodes de détection

Bien qu'il existe une certaine sensibilisation autour du phishing notamment par les sites du gouvernement Français [19] soulignant les réflexes à avoir, se voit être une méthode insuffisante depuis que les arnaques ont évolué, à travers les techniques et notamment dans l'ingénierie sociale. Pour se protéger de ces attaques, les logiciels et les experts se basent sur certains outils et technologies pour gagner en précision.

Dans les sous-sections suivantes, nous allons traiter les principales méthodes de détection.

Un ensemble des méthodes de détection générales sont présentées dans la Figure 1.

2.2.1 Détection basée sur les listes noires

Une méthode des plus classiques repose sur l'utilisation de bases de données connues, contenant des adresses mail, numéros de téléphone, des liens malveillants. Ces informations sont basées sur des cas de phishing ayant déjà eu lieu ou par d'autres informations diverses.

Des bases de données officielles sont retrouvables en ligne et mises à jour régulièrement par des organismes de sécurité, des services de renseignement sur les menaces. Un exemple avec Google Safe Browsing qui répertorie et met à jour ses listes noires pour les navigateurs Web et tous leurs services attachés à Google. [8] Uribl, un autre site recense les noms de domaine de sites hébergeant des virus, malwares et spywares [27]. Des listes sont disponibles sur certains sites : Infoserice [15]

Cette méthode est très utile et simple d'implémentation qui permet d'éviter une circulation encore plus massive du phishing. Cependant, comme il est précisé d'après des études Esposito [11], que cette méthode présente très rapidement des limites contre les nouveaux cas de phishing (appelé aussi le 0-Day Phishing¹), il est toujours possible aux attaquants de contourner ce mécanisme en créant de nouvelles adresses mails, de nouveaux noms de domaines, l'usurpation de numéros de téléphone.

2.2.2 Détection par règles heuristique

Beaucoup de techniques d'anti-phishing se basent sur des règles heuristiques. L'analyse par règle heuristique est une méthode de détection se basant sur des règles prédéfinies afin d'identifier des comportements suspects dans les e-mails. Contrairement aux listes noires, la détection de nouvelles attaques obtient un plus grand taux de détection. Cette méthode va chercher à identifier

des patterns anormaux ou des caractéristiques typiques d'attaques de phishing afin d'attribuer un score heuristique à chaque e-mail.

Ce score heuristique est attribué en fonction de plusieurs éléments comme :

- La présence de mots-clés suspects associés aux tentatives d'hameçonnage (ex. : "mise à jour de compte", "urgence", "mot de passe").
- La vérification des liens et des domaines, notamment en identifiant les domaines trompeurs ou les liens masqués.
- L'analyse des en-têtes d'e-mail, permettant de détecter les anomalies dans les informations d'expéditeur.
- La présence de fautes de frappe ou fautes grammaticales, souvent observées dans les e-mails frauduleux.
- L'utilisation de techniques d'obfuscation², comme le remplacement de caractères visuellement similaires

Contrairement à la précédente, cette méthode permet de s'adapter aux nouvelles attaques de phishing en identifiant des modèles de fraude même lorsqu'ils ne figurent pas encore dans les bases de données de menaces connues.

Néanmoins, d'après certains articles sur ScholarWorks [23], cette approche peut aussi faire paraître certaines limites, notamment dans un surplus de faux positifs répertoriés dans certains cas où ce surplus serait expliqué par des règles trop strictes.

Pour améliorer cette détection, l'utilisation de cette approche sur l'analyse heuristique est utilisée avec l'apprentissage automatique pour de meilleures performances.

[20]

2.2.3 Apprentissage automatique

L'apprentissage automatique (Machine Learning / ML) a démontré une amélioration en précision et performances supérieurs aux autres méthodes de détections. [3]

Ces modèles peuvent prédire en se basant sur des données existantes, des attaques nouvelles ou inconnues.

Pour le phishing, le ML utilise un jeu de données composé d'emails, SMS, des pages web considérés comme du phishing et d'autre de confiance, afin de reconnaître les différents patterns caractéristiques des attaques de phishing.

- Analyse du texte : présence de mots-clés suspects, ton alarmiste, fautes de grammaire et d'orthographe.

¹désigne une attaque de phishing exploitant une vulnérabilité inconnue ou récente, avant qu'une défense efficace ne soit mise en place

²désigne l'utilisation de méthodes visant à rendre difficile la détection ou l'analyse d'une information, souvent par la modification de son format ou de sa structure.

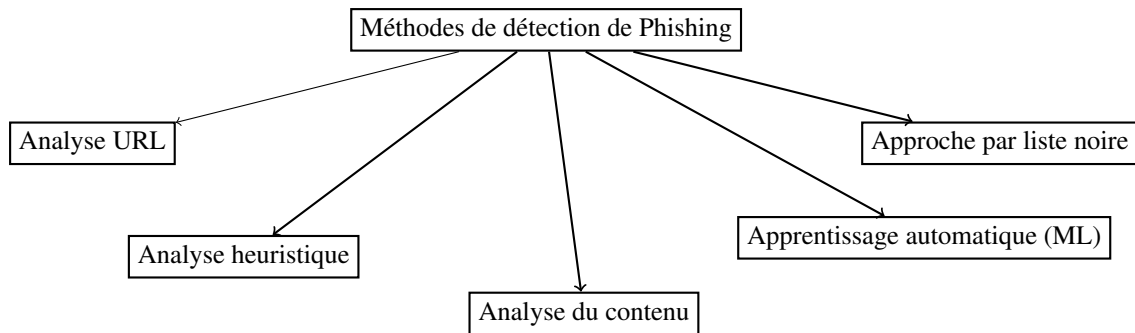


Figure 1: Méthodes d'identification de phishing existantes

Référence	Jeu de données utilisé (safe/phishing)	Méthode proposée	Précision (%)
Fette et al. [13]	6950 / 860	PILFER (LIBSVM - SVM public)	99.00
Abu-Nimeh et al. [1]	1700 / 1700	6 Classifieurs - LR, CART, BART, SVM, RF, NNet	95.11
Chandrasekaran et al. [7]	100 / 100	SVM à classe unique	95.00
Rathod and Patterwar [21]	2500 / 2100	Classificateur bayésien (Naïve Bayes)	96.46
Rawal et al. [22]	414 / 1605	Random Forest et SVM	99.87
Hota et al. [14]	Jeu de données public	RRFST avec C4.5 et CART	99.27
Mbah et al. [18]	6951 / 2357	KNN et Arbre de décision (J48)	93.11
Emilin Shyni et al. [10]	5260 / 0	Multi-classifieur - SVM, Random Forest, LogitBoost	96.30
Smadi et al. [24]	5000 / 5000	Algorithme de classification J48	98.11
Sonowal [25]	1604 / 1824	Sélection de caractéristiques par recherche binaire	97.41
Li et al. [17]	Jeu de données public	SVM avec AdaBoost	97.61
Jameel and George [16]	3000 / 3000	Réseau de neurones Feed Forward	98.72
Aljofey et al. [2]	Jeu de données public	Réseau de neurones récurrent convolutionnel	95.02
Fang et al. [12]	Combinaison de divers jeux de données publics	Modèle THEMIS basé sur les CNN	99.848
Bagui et al. [5]	14 950 / 3416	Réseau de neurones convolutionnel	95.97

Table 1: Performances des modèles d'apprentissage automatique. [9]

- Analyse des liens (URLs) : longueur, présence de caractères spéciaux, redirections suspectes, comparaison avec des domaines légitimes.
- Analyse des en-têtes d'email : incohérence entre l'expéditeur et le domaine du serveur.
- Analyse des pièces jointes : extensions suspectes, présence de scripts malveillants.

Selon l'étude de Dhruv Rathee [9], rassemblant plusieurs travaux antérieurs (voir Table 1), diverses méthodes d'apprentissage automatique ont été employées pour la détection des attaques de phishing. Ces approches incluent :

- SVM (Support Vector Machine)
- régression logistique
- arbres de décision
- réseaux de neurones
- forêts aléatoires
- Bayesian classifier
- k-Nearest Neighbors (k-NN)

Plusieurs de ces modèles ont démontré des taux de vrais positifs supérieurs à 95 %. Toutefois, malgré ces avancées significatives, des défis subsistent, notamment en matière de sélection des caractéristiques

pertinentes et d'adaptation aux nouvelles techniques d'attaque. Ces méthodes continuent d'évoluer, avec pour objectif d'améliorer davantage la précision et la robustesse des systèmes de détection. Un article scientifique [4] nous permet de vérifier une fois de plus ces statistiques et ses propos.

L'apprentissage nécessite un grand volume de données pour l'entraînement pouvant vite devenir coûteux en termes de calcul et de ressources. Des difficultés d'interprétation des décisions du modèle peuvent aussi poser problèmes. Ces désavantages sont expliqués dans l'article [4].

L'article conclut que cette amélioration considérable dans la détection du phishing, la sélection des caractéristiques et l'adaptation aux nouvelles techniques d'attaques restent des défis majeurs. Mais ces méthodes sont encore en voie d'évolution pour une précision plus prometteuse.

3 LLM

Les LLMs par leurs récentes performances sont de plus en plus sollicités, grâce à l'utilisation de modèles massifs et des réseaux neuronaux. Le but de l'expérimentation à venir étant de mettre à profit l'aspect globale des connaissances du LLM en utilisant sa capacité à comprendre des textes de manière efficace pour analyser les contenus des mails de phishing, y compris les subtilités linguistiques, les intentions cachées, le sens de la phrase et les techniques de manipulation

utilisées dans le principe de l'ingénierie sociale, ce qui les rend très adaptés aux problématiques rencontrées dans les détectons de phishing sophistiquées actuelles et futures. Permettant aussi d'utiliser une seule méthode pour toutes les langues possibles utilisées dans l'email potentiellement malveillant.

Comparé aux ML, le LLM peut détecter des patterns de phishing plus complexes qui ne sont pas nécessairement visibles dans les données d'entraînement en raison de leur compréhension linguistique du contenu, c'est-à-dire qu'elles seraient dans beaucoup de situations, être capables de s'adapter à des nouveaux types d'attaques de phishing inconnus.

3.1 classification

Les performances de modèle de détection pour notre étude sont mesurées par des métriques de classification :

- **Précision** (Precision) :

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

où TP représente le nombre de vrais positifs (phishing correctement identifié), et FP le nombre de faux positifs (e-mails légitimes identifiés à tort comme phishing).

4 Expérimentation / Simulation

La conception de ce service de détection de phishing par LLM découle d'une réflexion liée à notre pratique quotidienne en tant que développeur. L'utilisation fréquente de modèles avancés tels que ChatGPT, LLaMA ou Claude dans divers processus de développement a soulevé une question centrale : ces modèles, initialement conçus pour des tâches de génération et de compréhension du langage, pourraient-ils également être exploités pour remplacer des tâches complexes effectuées par des algorithmes de Machine Learning spécialisés ?

L'objectif de cette expérimentation est d'évaluer la faisabilité de substituer un modèle de machine learning classique, souvent coûteux en termes de ressources et de temps d'entraînement, par un service basé sur un LLM. Cette approche pourrait non seulement réduire les coûts liés à l'entraînement de modèles spécifiques, mais également accélérer le traitement, tout en conservant un niveau de performance satisfaisant pour la détection de phishing.

Pour ce faire, nous avons développé une application permettre à n'importe quel utilisateur de s'assurer de la nature malveillante d'un email gratuitement et en une dizaine de secondes. Ce service utilise l'API OpenAI pour analyser automatiquement le contenu de l'email et fournir un indice de probabilité indiquant s'il s'agit d'une tentative de phishing.

4.1 Spamurai

Ce service a pour rôle de faciliter notre étude sur l'état actuels des LLMs pour le phishing par mail. Mais aussi utilisable pour une utilisation personnelle.

Le déroulement du service est le suivant :

1. - L'utilisateur transfère un mail suspicieux à l'adresse **Spamurai.analysis@gmail.com**
2. extraction des features de l'email (adresse expéditeur, sujet, contenu)
3. Traitement et analyse du mail par nos serveurs via l'API OpenAI pour déterminer la dangerosité du mail en question
4. Envois du résultat de la requête à l'utilisateur sous forme de réponse à son e-mail

L'application est en ligne en permanence, pour aider n'importe qui ayant un doute sur l'authenticité d'un mail.

4.2 Protocole

4.2.1 Dataset

Pour évaluer les performances de notre approche, nous avons utilisé le jeu de données [6] disponible sur Kaggle. Ce jeu de données comprend 18 600 exemples d'emails répartis en trois colonnes :

1. L'indice du mail dans le dataset
2. Le corps du message (*body*), contenant le texte intégral de l'email ;
3. La catégorie de l'email, indiquant s'il s'agit d'un message légitime (Safe Email) ou d'une tentative de phishing (Phishing Email).

Dont :

- 61% d'emails considérés comme sûrs (Safe Email*) ;
- 39% d'emails identifiés comme des tentatives de phishing (Phishing Email).

4.2.2 Modèle

Dans cette étude, nous avons choisi d'utiliser le modèle 4o-mini d'OpenAI pour la détection de phishing. Ce modèle, bien que de taille réduite par rapport à d'autres modèles plus complexes, offre un bon compromis entre performance et coût avec un tarif de 0.15\$ / millions de tokens.

4.2.3 Prompt

Le rôle du prompt dans l'analyse est l'élément central dans l'efficacité de notre méthode. Il représente l'instruction donnée au modèle afin de guider son raisonnement et d'obtenir une analyse pertinente.

Pour cette étude, nous avons confectionné un prompt basé sur une série de critères essentiels permettant de détecter les tentatives de phishing :

Tout d'abord, le modèle est conditionné dès le début pour se comporter comme un expert en cybersécurité, spécialisé dans la détection de phishing, afin de maximiser la pertinence de son analyse.

- **Adresse de l'expéditeur** : Vérification des noms de domaine suspects, caractères inhabituels ou imitations de marques connues.
- **Objet du mail** : Identification des tactiques d'urgence ou de manipulation émotionnelle.
- **Contenu du message** : Détection des fautes grammaticales, demandes d'informations sensibles et liens suspects.
- **Techniques de manipulation psychologique** : Analyse des messages utilisant des figures d'autorité ou des menaces.
- **Urgence** : Présence de pressions pour obtenir une réponse rapide.

Le modèle retourne uniquement un score entre 0 et 100, indiquant la probabilité que l'email soit une tentative de phishing afin de pouvoir récupérer le résultat à chaque prompt sans risquer de ne pas pouvoir retrouver le score attribué dans la réponse du LLM.

Bien que des prompts plus complexes puissent fournir de meilleures performances, ils sont souvent moins fiables à grande échelle en raison de leur sensibilité accrue aux faux positifs, c'est pourquoi nous avons choisi une approche plus simplifiée mais plus robuste pour assurer une détection équilibrée.

4.2.4 Seuil de Détection

L'API utilisée retourne un score compris entre 0 et 100% indiquant la probabilité qu'un email soit du phishing. Afin d'optimiser la précision de la détection, une expérimentation a été menée pour déterminer un seuil approprié. Comme illustré dans la Figure 2, un taux de précision maximal a été observé lorsque le seuil de classification des emails comme étant du phishing a été fixé à 75. Ce seuil permet d'atteindre un équilibre optimal entre la détection correcte des emails de phishing et la minimisation des erreurs de classification.

5 Resultats

Nous avons évalué les performances de notre approche sur un échantillon de 51 emails tirés du jeu de données, 32 mails non fishing et 19 mails fishing.

Bien que cet échantillon soit relativement restreint par rapport à ceux utilisés dans les approches basées sur des modèles d'apprentissage automatique traditionnels, il convient de souligner que, du fait de la nature générale et robuste de notre modèle, un tel échantillon reste adéquat.

Étant donné que le modèle est conçu pour gérer un large éventail de cas sans nécessiter de cas spécifiques pour son apprentissage, les résultats obtenus ne dépendent

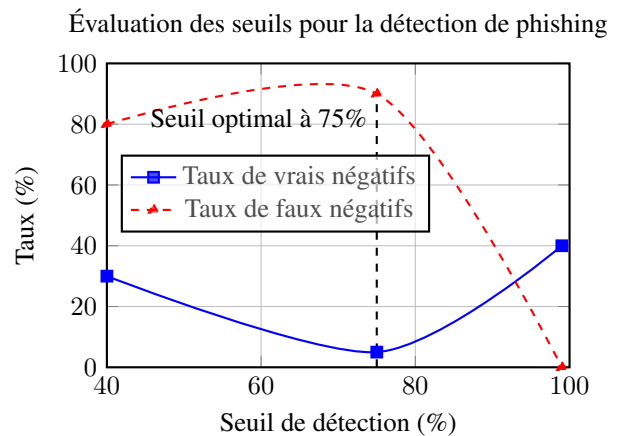


Figure 2: Impact des différents seuils sur les taux de détection. Un compromis optimal est atteint à un seuil de 75%, offrant une précision finale de 90.19%.

pas fortement de la taille de l'échantillon. De plus, en raison des contraintes liées à l'usage pratique d'un service de détection par email, il s'avère difficile de tester sur un ensemble plus vaste.

	Prédis: Safe	Prédis: Phishing	Total
Safe	27	5	32
Phishing	0	19	19

5.1 Analyse des résultats

5.1.1 Faux positifs

Pour les mails phishing on a observé durant toute notre expérimentation très peu de cas de faux positifs (uniquement quand le seuil était > 95).

Pour les mails non phishing identifié une limite dans l'approche basée LLM due à la censure du modèle. En effet, certains messages ont été mal classés comme phishing en raison de facteurs internes liés à la censure du modèle, où le raisonnement du modèle a été entravé par des restrictions imposées pendant l'apprentissage.

5.1.2 Vrais positifs

En ce qui concerne les mails de phishing, notre modèle a montré un excellent taux de détection des facteurs caractéristiques associés au phishing. Les éléments souvent utilisés par les cybercriminels pour manipuler les victimes ont été efficacement identifiés, ce qui a conduit à une classification correcte de la majorité des emails malveillants.

Pour les emails légitimes, en dehors des problèmes précédemment mentionnés (relatifs aux faux positifs), le modèle a montré une grande précision dans la reconnaissance des emails sûrs, avec très peu de cas de classification erronée. Ces résultats suggèrent que le modèle possède une capacité robuste à distinguer les mails légitimes des tentatives de phishing dans la plupart des cas.

6 Limitations

Bien que les résultats soient prometteurs, certaines limitations sont à considérer. Par exemple, la taille restreinte du dataset testée (51 emails) pourrait ne pas refléter la diversité complète des cas de phishing présents dans des environnements réels. Pour cela il aurait fallu créer un service permettant une plus simple transition du dataset à l'envoi de mail par un autre outil à développer nous même.

Également l'utilisation des modèles plus coûteux et plus efficace garantirai une meilleure précision de notre recherche en prenant aussi avantage des modèles multimodaux permettant de fournir au modèles d'éventuelles pièces-jointes incluses dans le mail (non traitées actuellement). Enfin on remarque que le biais habituel de la surcouche des modèles empêche le raisonnement sur les emails contenant du contenu grossier, diffamatoire, à caractère sexuel et catégorise instantanément le mail comme contenu phishing.

7 Conclusion

Dans cette étude, nous avons pu expérimenter l'efficacité des LLM pour la détection de phishing. En comparant les résultats aux méthodes plus traditionnelles, on obtient des performances supérieures notamment en termes de capacité à s'adapter aux nouvelles attaques, grâce à leur flexibilité et leur compréhension sémantique des emails.

Cependant, en comparant avec les résultats d'études de détection en Machine learning, on obtient une précision qui est en moyenne inférieure à celles-ci. Ce décalage peut être dû à une imprécision du modèle choisi. Malgré ces résultats, les LLMs dans le temps se verront sûrement augmenter en précision car ils possèdent un fort potentiel pour dépasser les limites actuelles des autres méthodes, que ce soit dans la détection du phishing ou certains domaines très complexes.

References

- [1] S. Abu-Nimeh, D. Nappa, X. Wang, and S. Nair. 2007. Comparison of machine learning techniques for phishing detection. In *APWG eCrime Researchers Summit*, Pittsburgh, USA.
- [2] A. Aljofey, Q. Jiang, Q. Qu, M. Huang, and J.-P. Niyigena. 2020. An effective phishing detection model based on character level convolutional neural network from url. *Electronics*, 9(9):1514.
- [3] Eman Abdelfattah Ammar Odeh, Ismail Keshta. 2020. [Machine learning techniques for detection of website phishing: A review for promises and challenges](#). *IEEE Xplore*.
- [4] Anu Vazhayil, Harikrishnan NB, Vinayakumar R, Soman KP. 2020. [Phishing email detection using classical machine learning techniques](#). *Amrita School of Engineering*.
- [5] S. Bagui, D. Nandi, and R. J. White. 2021. Machine learning and deep learning for phishing e-mail classification using one-hot encoding. *Journal of Computer Science*, 17(7):610–623.
- [6] Subhadeep Chakraborty. 2022. [Phishing emails dataset](#). Accessed: January 23, 2025.
- [7] M. Chandrasekaran, K. Narayanan, and S. Upadhyaya. 2006. Phishing e-mail detection based on structural properties. In *First Annual Symposium on Information Assurance: Intrusion Detection and Prevention*, pages 2–8, New York.
- [8] Wikipedia Contributors. 2025. [Google safe browsing](#).
- [9] Suman Mann Dhruv Rathee. 2022. [Detection of e-mail phishing attacks – using machine learning and deep learning](#). *International Journal of Computer Applications*. Accessed: January 23, 2025.
- [10] C. Emilin Shyni, S. Sarju, and S. Swamynathan. 2016. A multi-classifier based prediction model for phishing emails detection using topic modelling, named entity recognition and image processing. *Circuits and Systems*, 7:2507–2520.
- [11] Andrea et d’autres auteurs Esposito. 2021. [An effective detection approach for phishing websites using url and html features](#). *PubMed Central*.
- [12] Y. Fang, C. Zhang, C. Huang, L. Liu, and Y. Yang. 2019. Phishing e-mail detection using improved rcnn model with multilevel vectors and attention mechanism. *IEEE Access*, 7:56329–56340.
- [13] I. Fette, N. Sadeh, and A. Tomasic. 2006. Learning to detect phishing e-mails. Technical report, Institute of Software Research International, School of Computer Science, Carnegie Mellon University.
- [14] H. Hota, A.K. Shrivastava, and Rahul Hota. 2018. An ensemble model for detecting phishing attack with proposed remove-replace feature selection technique. *Computer Science*, 132:900–907.
- [15] ABE Infoservice. 2025. [Listes noires des autorités](#).
- [16] Noor M. Jameel and Loay George. 2013. Detection of phishing e-mails using feed forward neural network. *International Journal of Computer Applications*.
- [17] Y. Li, Z. Yang, X. Chen, H. Yuan, and W. Liu. 2019. A stacking model using url and html features for phishing webpage detection. *Future Generation Computer Systems*, 94:27–39.
- [18] K. F. Mbah, A. H. Lashkari, and A. A. Ghorbani. 2022. A phishing e-mail detection approach using machine learning techniques. *World Academy of Science, Engineering and Technology, Computer and Information Engineering*, 3:2333.
- [19] Ministère de l’Économie, des Finances et de la Souveraineté industrielle et numérique. [Phishing, hameçonnage et filoutage](#).
- [20] Hong Qin, Ramprasath Jayaprakash, and Saurav et d’autres auteurs Mallik. 2021. [Heuristic machine learning approaches for identifying phishing threats across web and email platforms](#). *PubMed Central*.
- [21] Sunil B. Rathod and Tareek M. Pattewar. 2015. Content based spam detection in e-mail using bayesian classifier. In *IEEE ICCSP Conference*.
- [22] Srishti Rawal, Bhuvan Rawal, Aakhila Shaheen, and Shubham Malik. 2017. Phishing detection in e-mails using machine learning. *International Journal of Applied Information Systems*, 12:21–24.
- [23] California State University ScholarWorks. 2025. [Research paper on phishing detection techniques](#). *California State University ScholarWorks*.
- [24] Sami Smadi, Nauman Aslam, Li Zhang, Rafe Alasem, and M.A. Hossain. 2015. Detection of phishing e-mails using data mining algorithms. In *9th International Conference on Software, Knowledge, Information Management and Applications (SKIMA)*.
- [25] G. Sonowal. 2020. Phishing e-mail detection based on binary search feature selection. *SN Computer Science*, 1.
- [26] Statista Research Department. 2024. [Number of cyberattacks worldwide per year](#).
- [27] URIBL. 2025. [Uribl - universal resource identifier blacklist](#).