Rédigé par **Joshua LEMOINE, Berachem MARKRIA, Abdallah M'CHIRI,**

Alessandro VILLA, Ismaël MOSTESFA-SBA

Prédiction des cyber-attaques

Machine Learning

Tables des matières

1. Introduction

1.1 Contexte et problématique

Dans un monde de plus en plus connecté, les cyberattaques posent une menace croissante à la sécurité des systèmes d'information, rendant difficile la surveillance efficace des réseaux complexes. Le **machine learning** (ML) se révèle essentiel pour **détecter** des anomalies dans de vastes volumes de données, permettant ainsi **d'anticiper** les attaques plus rapidement que les méthodes classiques. Cependant, la complexité croissante des réseaux rend ce processus toujours plus exigeant, d'où la nécessité de solutions de défense sophistiquées et adaptées.

1.2 Objectifs du rapport

Ce rapport vise à explorer l'application du machine learning dans la **prédiction** des cyberattaques, en comparant les techniques supervisées, non supervisées et semi-supervisées. Il identifiera les solutions les plus efficaces en fonction des contraintes techniques et des limites des ensembles de données. Des recommandations seront proposées pour garantir la robustesse des approches tout en assurant la confidentialité des données.

2. Contexte de projet

1.1 Présentation du projet

Ce projet vise à **prédire** les cyberattaques à l'aide du **machine learning**, un enjeu crucial face à l'augmentation des **menaces** informatiques. En exploitant des données réseau pour **identifier** des comportements suspects, le modèle développé permettra de réagir rapidement aux menaces internes et externes. L'objectif est de mettre en place un système capable de prévenir les intrusions avant qu'elles ne causent des dommages significatifs.

1.2 Enjeux et importance de la prédiction des cyberattaques en entreprises

La prédiction proactive des cyberattaques est devenue un impératif stratégique pour les entreprises, notamment en raison des pertes financières qu'elles peuvent engendrer. En 2023, le coût moyen d'une violation de données a atteint **4,45 millions de dollars**, une augmentation de **15 %** en trois ans *Stormshield*. De plus, les cyberattaques ont augmenté de **104 %** en 2023 *Business Wire*, ce qui souligne la nécessité d'améliorer les capacités de détection et de réaction face à ces menaces. Le recours au **machine learning** permet d'analyser les données en temps réel, augmentant ainsi la précision et la rapidité d'identification des comportements anormaux, et renforçant la protection des systèmes.

Les principaux enjeux incluent la protection des données sensibles, l'intégrité des systèmes, la disponibilité des services, et la conformité aux réglementations. La complexité croissante des menaces exige des solutions adaptatives. Un système prédictif basé sur le **machine learning** permet d'améliorer la détection tout en **réduisant** les besoins en ressources humaines.

"La cybersécurité n'est pas un problème technologique, c'est un problème économique, social et politique," (Bruce Schneier)

3. Recherche de Datasets

3.1 Jeux de données disponibles pour la cybersécurité

Pour entraîner et évaluer les modèles de machine learning dans le domaine de la cybersécurité, plusieurs jeux de données sont disponibles. Ces jeux de données contiennent généralement des enregistrements de trafic réseau simulant diverses attaques, ainsi que des événements liés à des comportements normaux. Parmi les jeux de données les plus utilisés, nous pouvons citer :

3.1.1 UNSW-NB15

Le jeu de données UNSW-NB15 est l'un des jeux de données modernes les plus utilisés pour la détection d'intrusions dans le domaine de la cybersécurité. Il a été développé par l'Australian Cyber Security Centre (ACSC) en 2015 dans le but de surmonter les limitations des anciens jeux de données comme KDD99. UNSW-NB15 capture des scénarios de trafic réseau réel ainsi que des attaques récentes. Ce jeu de données comprend 49 caractéristiques décrivant le flux réseau, telles que la durée de la connexion, le type de service, et les taux de transfert, ainsi que neuf types d'attaques, notamment les attaques par injection SQL, les attaques DoS (déni de service), les attaques par malware et les scans réseau.

Ce dataset est structuré pour offrir un environnement de test plus réaliste, combinant à la fois des comportements légitimes et des attaques réelles. En raison de son approche équilibrée et plus actuelle, il est particulièrement adapté pour la formation des modèles de machine learning dans des contextes modernes de cybersécurité. Il est souvent considéré comme un bon compromis entre la complexité des données et la diversité des menaces qu'il représente.

L'UNSW-NB15 inclut également un ensemble de données équilibré, permettant d'éviter les problèmes de déséquilibre des classes fréquemment rencontrés dans d'autres

datasets, ce qui le rend plus efficace pour évaluer les performances des modèles sur une large gamme d'attaques.

3.1.2 KDD99

Le jeu de données KDD99 (KDD Cup 1999) est un ensemble de données historique dans la recherche en cybersécurité, conçu pour une compétition sur la détection d'intrusions en réseau. Il contient des enregistrements de connexion réseau, avec des catégories comme des connexions normales ou diverses formes d'attaques. Les attaques sont réparties en quatre grandes catégories : les attaques par déni de service (DoS), les accès non autorisés (U2R - User to Root), les attaques à distance (R2L - Remote to Local), et les scans de ports ou de réseaux.

Le dataset se compose de 41 caractéristiques extraites de chaque connexion réseau, et inclut un large volume de données (environ 4 millions d'enregistrements). KDD99 est particulièrement apprécié pour sa taille, qui permet aux chercheurs de développer des modèles robustes, bien qu'il soit reconnu pour certaines lacunes, telles que la présence de données redondantes et des scénarios d'attaques souvent simplifiés par rapport à la réalité.

Malgré ces limites, KDD99 reste largement utilisé dans les recherches pour la détection d'intrusions, notamment pour l'entraînement initial de modèles et pour la comparaison avec des approches plus modernes. Toutefois, il est généralement recommandé de compléter ce dataset avec d'autres jeux de données plus récents pour tenir compte des menaces actuelles, qui évoluent rapidement et sont souvent plus complexes que celles représentées dans KDD99.

3.2 Impact et Limites des Données sur les Résultats des Modèles

		KDD99 data set		UNSW-NB15 data set	
Techniques	Reference	Accuracy (%)	FAR (%)	Accuracy (%)	FAR (%)
DT	(Bro-IDS Tool, 2014)	92.30	11.71	85.56	15.78
LR	(Witten & Mining, 2005)	92.75	-	83.15	18.48
NB	(Shyu et al., 2005)	95	5	82.07	18.56
ANN	(Witten & Mining, 2005)	97.04	1.48	81.34	21.13
EM clustering	(Salem & Buehler, 2012)	78.06	10.37	78.47	23.79

Source Tableau

Les performances des modèles de détection d'intrusions sont fortement influencées par la qualité et la complexité des jeux de données utilisés pour leur entraînement et leur évaluation. Deux ensembles de données largement utilisés dans le domaine de la sécurité des réseaux sont le KDD99 et l'UNSW-NB15, chacun présentant des

particularités qui influencent les résultats obtenus par différentes techniques de classification.

Le jeu de données KDD99 est un ancien jeu de données, dont les comportements de trafic réseau sont moins représentatifs des attaques modernes. Cela explique pourquoi les modèles entraînés avec ce jeu obtiennent de meilleurs résultats en termes d'exactitude (accuracy) et de taux de fausses alertes (FAR), car les attaques et les comportements normaux qu'il contient sont plus facilement distinguables. Par exemple, dans le tableau comparatif, on constate que pour la technique des arbres de décision (DT), le KDD99 atteint une exactitude de 92,30 % avec un FAR de 11,71 %, tandis que d'autres techniques comme les réseaux bayésiens (NB) et les réseaux neuronaux (ANN) offrent respectivement des performances encore plus élevées avec 95 % et 97,04 % d'exactitude, et des FAR réduits à 5 % et 1,48 %.

À l'inverse, le jeu de données UNSW-NB15 est plus récent et inclut un large éventail de menaces contemporaines et de comportements réseau. Cela rend la distinction entre le trafic normal et les attaques plus difficile, augmentant ainsi la complexité du jeu de données. Le réseau UNSW-NB15 montre des performances globalement inférieures pour les mêmes techniques. Par exemple, pour les arbres de décision (DT), l'exactitude tombe à 85,56 % avec un FAR plus élevé de 15,78 %. D'autres techniques comme les réseaux bayésiens et les réseaux neuronaux subissent également une baisse de performances, avec des exactitudes respectives de 82,07 % et 81,34 %, et des FAR augmentés (18,56 % et 21,13 %).

Les résultats inférieurs sur l'UNSW-NB15 sont dus à plusieurs facteurs. D'une part, les similitudes dans les caractéristiques des observations normales et des attaques compliquent la tâche des modèles. En effet, les tests statistiques montrent que les distributions des ensembles d'entraînement et de test de l'UNSW-NB15 sont non linéaires et non-normales, avec des corrélations élevées entre les caractéristiques, ce qui augmente la complexité de la classification. D'autre part, la présence de comportements réseau modernes rend plus difficile la détection des anomalies par les techniques classiques, car celles-ci ne sont pas toujours adaptées aux nouveaux types de menaces.

En résumé, bien que le jeu de données KDD99 permette d'obtenir de meilleures performances globales, il est moins représentatif des défis actuels en matière de sécurité des réseaux. En revanche, l'UNSW-NB15, malgré des résultats plus modestes, offre un cadre plus réaliste pour évaluer l'efficacité des systèmes de détection d'intrusions face aux menaces modernes. Cela met en lumière l'importance de la qualité et de la mise à jour des jeux de données pour garantir la pertinence des modèles face aux attaques contemporaines.

4. Types d’Apprentissage

Dans le cadre de la création d’une IA basée sur le machine learning il existe plusieurs manières d'entraîner notre modèle. Chaque approche offre des avantages et des inconvénients spécifiques en fonction des scénarios d'application de la nature des

menaces, la disponibilité des données étiquetées. Nous discuterons ici des trois types d'apprentissage les plus répandus : supervisé, semi-supervisé et non-supervisé

4.1 Apprentissage supervisé

L'apprentissage supervisé est une méthode dans laquelle un modèle est formé sur un ensemble de données étiquetées, où chaque exemple de formation comprend des entrées et une sortie attendue. Le but du modèle est d'apprendre la relation entre les entrées et les sorties, de façon à pouvoir prédire l'étiquette correcte pour de nouvelles données non vues.

Ce type d'apprentissage étant très dépendant du jeu de données, il est très efficace pour des tâches ou les données sont disponibles en abondance, correctement étiquetées et peu sujettes à l'évolution au cours du temps. Par exemple, la prévision du nombre de calories d'un aliment en fonction de la présence de certains composés (les lipides auront toujours le même impact sur les calories dans 10 ans).

Dans le cas de la cybersécurité, il existe un grand nombre de données correctement étiquetées. Cependant, comme vu dans la partie *4.2 Impact et Limites des Données Sur les Résultats des Modèles* les attaques informatiques ont une grande tendance à évoluer au cours des années. Il est donc possible qu'un modèle entraîné par apprentissage supervisé ne soit valable que sur une période de quelques années. De plus, il faudrait attendre l'émergence de nouveaux datasets à jour pour réentrainer un nouveau modèle efficace.

4.2 Apprentissage non supervisé

A l'opposé de l'apprentissage supervisé l'apprentissage non supervisé utilise des données non étiquetées pour découvrir des structures cachées ou des relations dans les données. Il n'y a pas d'étiquette de sortie fournie, et le modèle doit identifier les motifs de manière autonome. Le modèle cherche à organiser les données en cluster ou à détecter des anomalies. Les algorithmes populaires incluent les k-means pour le clustering et les méthodes basées sur la densité pour la détection d'anomalies.

Ce type d'apprentissage à l'avantage de détecter facilement des menaces émergentes ou inconnues. Mais aussi de pouvoir s'entraîner sur un grand nombre de données non-étiquetées ce qui lui permettrait d'être mis à jour facilement au cours du temps.

Les modèles non supervisés peuvent produire un grand nombre de faux positifs car il est difficile de distinguer les anomalies légitimes des menaces. Les résultats peuvent également être difficiles à interpréter. Cela peut donc poser un réel problème d'efficacité comme discuté dans la partie *5.2 Limites et Impact des faux positifs et faux négatifs.*

4.3 Apprentissage semi-supervisé

Enfin, le troisième type d'apprentissage étudié ici est l'Apprentissage semi-supervisé. Il combine les principes de l'apprentissage supervisé et non supervisé, en utilisant à la fois des données étiquetées et non étiquetées. Le modèle utilise les données étiquetées pour guider l'apprentissage et exploite les données non étiquetées pour renforcer les motifs identifiés.

Il semble à première vue réunir le meilleur des deux mondes en permettant de générer un modèle actuellement fiable grâce au grand nombre de données étiquetées existant actuellement tout en permettant l'utilisation de données non étiquetées pour affiner la détection de motifs nouveaux au cours du temps.

Cependant ce type d'entraînement peut être très dur à mettre en place car il peut utiliser en réalité un grand nombre de techniques. Comme par exemple :

- ***l'apprentissage par transfert*** **:** Utilisation d'un modèle pré-entraîné sur une tâche pour accélérer l'entraînement sur une nouvelle tâche similaire. Cela suppose donc que l'on possède déjà un modèle efficace pour une tâche similaire.

- ***l'apprentissage par pseudo étiquetage*** **:** Génération des étiquettes artificielles pour des données non étiquetées en utilisant les prédictions d'un autre modèle pour les inclure dans l'entraînement. C'est un peu le serpent qui se mord la queue, besoin d'un model pour entraîner un autre model pour entraîner un autre model…

5. Étude des Métriques de Performance

5.1 Métriques de classification : Précision, Rappel, F1-score

Dans le domaine du Machine Learning, les métriques de classification jouent un rôle important dans l'évaluation de l'efficacité des systèmes de détection automatisée. Dans le contexte de la cybersécurité plusieurs de ces métriques sont particulièrement importantes.

Tout d'abord, la **précision** est utile pour mesurer la proportion d'alertes justifiées par le modèle. Par exemple, dans un système de détection de logiciels malveillants, une haute précision signifie que la majorité des fichiers identifiés comme malveillants sont réellement des menaces. Cependant, la précision seule n'est pas suffisante. Les attaques sont rares par rapport au trafic légitime. Dans un tel contexte, un modèle pourrait obtenir une précision élevée simplement en classant la majorité des éléments comme légitimes.

Il faut donc aussi considérer le **rappel** car il mesure la capacité du modèle à détecter toutes les menaces réelles, ce qui est essentiel pour minimiser les risques de failles de sécurité. Un rappel élevé indique que peu d'attaques passent inaperçues, ce qui est crucial dans les environnements critiques où les faux négatifs (menaces non détectées) peuvent avoir des conséquences graves.

Le **F1-score**, qui combine précision et rappel, est souvent considéré comme une métrique plus équilibrée dans le domaine de la cybersécurité. Il permet de prendre en compte à la fois les faux positifs (alertes incorrectes) et les faux négatifs (menaces non détectées). Un bon F1-score est essentiel pour s'assurer que le système de détection maintient un compromis optimal entre ces deux aspects, surtout lorsque les données sont déséquilibrées, avec plus de trafic légitime que d'attaques.

Dans certains cas, d'autres métriques peuvent être plus adaptées. Par exemple, la **matrice de confusion**, qui montre la répartition des vrais positifs, faux positifs, vrais négatifs et faux négatifs, fournit également une vue détaillée des erreurs du modèle et permet d'ajuster ses paramètres pour optimiser les performances selon les priorités spécifiques de sécurité.

5.2 Limites et Impact des faux positifs et faux négatifs

Les faux positifs et les faux négatifs constituent deux types d'erreurs dans les systèmes de classification, et leur gestion est particulièrement critique dans les applications de cybersécurité.

Premièrement, les **faux positifs** se produisent lorsqu'un modèle classe incorrectement une activité légitime comme une menace. Dans le contexte de la cybersécurité, cela peut entraîner des alertes inutiles. Un taux élevé de faux positifs peut non seulement surcharger les analystes en générant un volume élevé d'alertes à examiner, mais aussi diminuer leur confiance dans le système de détection, les incitant éventuellement à ignorer certaines alertes ou à désactiver des fonctions de sécurité critiques. Cette situation est particulièrement problématique dans les environnements à forte activité.

Les **faux négatifs**, eux, se produisent lorsqu'une menace réelle n'est pas détectée par le modèle. Ces erreurs sont généralement plus dangereuses que les faux positifs dans le domaine de la cybersécurité, car elles signifient que des attaques passent inaperçues et peuvent causer des dommages considérables. Le coût des faux négatifs peut être élevé, notamment lorsque les attaques réussissent à se propager avant d'être détectées, entraînant des pertes financières et une atteinte à la réputation de l'organisation.

En somme, dans le pire des cas, les **faux positifs** peuvent générer une perte de productivité ainsi que la création de nouvelle failles dues à la fatigue des analystes. Par contre les **faux positifs** peuvent avoir des conséquences bien plus graves, qu'elles soient financières ou techniques. C'est pour cela que les paramètres décrits dans la partie précédente sont si important dans un contexte de cybersécurité.

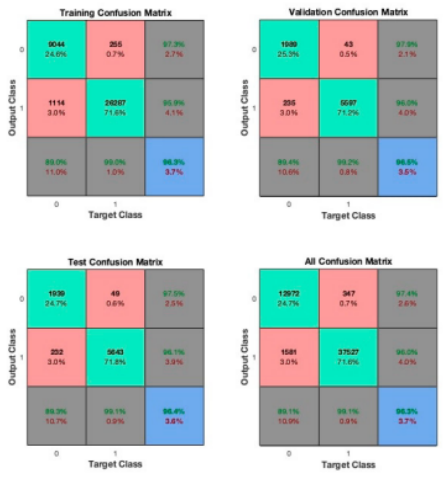
6. État de l'Art

6.1 Définition de l'état de l'art

L'**état de l'art** fait référence à une revue systématique des connaissances actuelles dans un domaine spécifique, permettant d'identifier les méthodes, techniques, et résultats les plus récents. Dans le contexte de la cybersécurité, il s'agit d'analyser les approches existantes en matière de détection et de prédiction des cyberattaques, en évaluant leur efficacité et leur pertinence face aux menaces contemporaines. Cela implique également une compréhension des défis et des lacunes des méthodes utilisées, ainsi que des directions futures de recherche.

6.2 Comparaison des modèles, limites et défis

Les **réseaux de neurones profonds** (DNN) se sont révélés particulièrement performants dans la détection des anomalies en cybersécurité, atteignant des taux de précision moyens entre 96% et 97% sur des datasets comme KDD99. En revanche, les forêts aléatoires et les SVM offrent une précision comparable tout en étant plus légers en termes de ressources computationnelles, mais ils peuvent montrer des limites lorsque les données sont massives ou déséquilibrées .



Les matrices du modèle DNN appliqué au dataset KDD99 montrent que[1] :

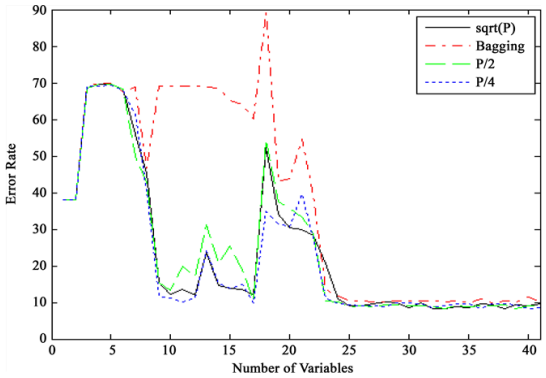
- **Précision** : Le DNN atteint jusqu'à 97.4% en phase de validation, tandis que les autres modèles montrent des performances légèrement inférieures
- **Sensibilité (Recall)** : Le DNN dans la détection des attaques avec un taux de rappel supérieur à 96% pour toutes les phases (entraînement, validation, test)
- **Spécificité (Precision) :** Les faux positifs sont inférieurs à 3%, ce qui est un indicateur important pour la fiabilité des détections

Les **forêts aléatoires** (Random Forest) sont également largement utilisées pour la détection des anomalies en cybersécurité. Sur le dataset **KDD99**, le modèle Random Forest offre une précision notable, bien que légèrement inférieure à celle des réseaux de neurones profonds (DNN). Cependant, il présente l'avantage d'une utilisation plus efficace des

[1] Source DNN ici

ressources computationnelles, notamment lorsque le nombre de variables est réduit, ce qui en fait une option viable pour les environnements à ressources limitées.

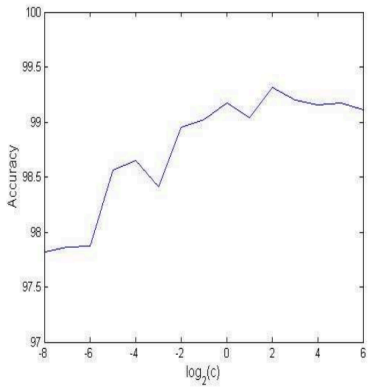
Les résultats du modèle Random Forest appliqué au dataset **KDD99** montrent que[2] :



- **Précision** : Le modèle RF avec 25 variables atteint une précision de **91.9%** en phase de test, contre **91.41%** pour le modèle avec 41 variables. Cette réduction du nombre de variables permet également de réduire le temps d'entraînement (7.98 minutes contre 10.62 minutes pour le modèle avec 41 variables).
- **Sensibilité (Recall)** : Le modèle Random Forest détecte efficacement les attaques **Dos** et **Normal**, avec des taux de rappel atteignant respectivement **98.94%** et **91.88%** pour le modèle à 25 variables.
- **Spécificité (Precision)** : La précision du modèle RF varie en fonction des types d'attaques détectées. Par exemple, pour les attaques **Dos**, le taux de précision atteint **98.94%**, tandis que pour les attaques **Probing**, il est plus faible à **55.53%**. Cependant, le modèle RF avec 25 variables a un taux de faux positifs plus bas que le modèle à 41 variables, notamment sur les attaques **Dos** (**5.82%** contre **7.52%**).

Les **machines à vecteurs de support** (SVM) ont également été testées sur le dataset KDD99, en utilisant différents noyaux : linéaire, polynomiale et radial basis (RBF). Les résultats montrent que le noyau radial basis (RBF) se démarque avec une précision de 92.99% sur les données de test, surpassant les noyaux linéaire (36.93%) et polynomiale (91.27%).

Les matrices de confusion pour le modèle SVM appliqué au dataset KDD99 montrent que[3] :



- **Précision** : Le noyau RBF atteint une précision de 96.32% pour la détection des attaques DoS, tandis que la précision moyenne des autres types d'attaques varie entre 57.36% (Probing) et 85.64% (R2L).
- **Sensibilité (Recall)** : Le noyau linéaire montre une performance limitée, particulièrement pour la détection des attaques R2L (44.07%) et U2R (2.12%), tandis que le noyau RBF améliore légèrement les résultats avec des taux de rappel de 85.64% pour R2L et 34.48% pour U2R.
- **Spécificité (Precision)** : Le noyau RBF offre une précision moyenne de 70%, avec des performances notables

[2] Source RF ici

[3] Source SVM ici

sur les attaques DoS (96.16%), mais des résultats plus faibles pour les attaques U2R (14.29%).

Ainsi on peut voir que le **DNN** est le modèle qui offre la meilleure précision globale, mais il demande des ressources computationnelles importantes. Il est idéal pour des systèmes où la performance est prioritaire et où les ressources sont abondantes.

D'autre part, la **Random Forest** est un excellent compromis en termes de précision et d'efficacité computationnelle, surtout pour des environnements où les ressources sont limitées, bien que sa précision soit plus faible pour certaines attaques spécifiques comme Probing.

Enfin, **SVM** avec noyau **RBF** offre une alternative compétitive avec une bonne précision pour certaines catégories d'attaques (comme DoS), mais sa performance peut varier fortement en fonction du noyau utilisé et des types d'attaques à détecter.

En fonction des besoins spécifiques en termes de rapidité, ressources disponibles et types d'attaques à détecter, chaque modèle présente des avantages et des limites qui le rendent plus ou moins adapté.

7. Considérations sur la Vie Privée et la Sécurité

7.1 Sensibilité des données dans les modèles de prédictions

L'un des enjeux majeurs dans la prédiction des cyberattaques est la gestion des données **sensibles** utilisées pour entraîner les modèles de machine learning. Les données collectées à partir des systèmes d'information et des réseaux peuvent contenir des informations **confidentielles**, dont la compromission pourrait engendrer des **violations** de la vie privée. Par conséquent, il est essentiel de mettre en place des mécanismes de protection, tels que le chiffrement des données et la pseudonymisation, pour garantir leur sécurité tout au long du processus de modélisation. Tel que **AES, RSA, Hachage…**

7.2 Régulations et conformité (ex. RGPD)

Les régulations comme le **RGPD** (Règlement Général sur la Protection des Données) imposent des **obligations** strictes concernant l'utilisation, le stockage et le partage des données personnelles. Dans le cadre de la **prédiction** des cyberattaques, les entreprises doivent s'assurer que leurs systèmes respectent ces exigences, notamment en **limitant** l'accès aux données personnelles, en obtenant les consentements nécessaires et en assurant la transparence quant à l'utilisation des données.

7.3 Sécurisation des données et solutions techniques

Pour renforcer la sécurité des données utilisées dans les modèles de machine learning, plusieurs solutions techniques peuvent être mises en œuvre. Parmi elles, on trouve le chiffrement des données en transit (**SSL, TLS…**) et au repos (**Bitlocker**), l'utilisation de réseaux privés virtuels (**VPN**) pour sécuriser les communications, ainsi que l'application de techniques de "differential privacy" qui permettent d'entraîner les modèles tout en minimisant les risques de ré-identification des utilisateurs. De plus, les systèmes doivent être régulièrement mis à jour pour contrer les nouvelles menaces et éviter les vulnérabilités exploitées par les cybercriminels.

8. Conclusion

L'apprentissage **supervisé**, bien qu'efficace dans les contextes où les données étiquetées sont disponibles, peut manquer de flexibilité face aux menaces en constante évolution. À l'inverse, l'apprentissage **non supervisé** est plus adapté pour détecter des anomalies ou des attaques émergentes, mais il peut engendrer un nombre élevé de faux positifs, ce qui complique sa gestion au quotidien. L'approche **semi-supervisée** représente une alternative intéressante, combinant le meilleur des deux mondes, en permettant l'utilisation des données étiquetées et non étiquetées. Toutefois, sa mise en œuvre nécessite une gestion technique complexe et un ajustement permanent.

La sélection des datasets est un autre élément crucial pour garantir l'efficacité du modèle. Bien que des ensembles de données historiques comme **KDD99** offrent de bonnes bases de comparaison, ils ne sont plus suffisamment représentatifs des menaces actuelles. En ce sens, des datasets plus récents comme **UNSW-NB15**, qui reflètent mieux les comportements réseau contemporains, sont fortement recommandés pour former des modèles plus robustes et adaptés aux environnements actuels.

En conclusion, le choix du modèle et de l'approche dépendra des besoins spécifiques de l'entreprise, de la nature des menaces à anticiper et des ressources disponibles. Un modèle **semi-supervisé** pourrait être une solution flexible et évolutive pour certaines organisations, tandis que d'autres pourraient opter pour un apprentissage **supervisé** ou non supervisé en fonction de leurs priorités et des données à disposition. Il est donc essentiel d'adopter une approche modulaire et de rester adaptable pour suivre l'évolution rapide des cybermenaces et optimiser la protection des systèmes d'information.

Sitographie

The UNSW-NB15 Dataset | UNSW Research. (s. d.). https://research.unsw.edu.au/projects/unsw-nb15-dataset

Bibliographie

Almajed, R., Ibrahim, A., Abualkishik, A. Z., Mourad, N., & Almansour, F. A. (2022). Using machine learning algorithm for detection of cyber-attacks in cyber physical systems. *Periodicals Of Engineering And Natural Sciences (PEN)*, *10*(3), 261. https://doi.org/10.21533/pen.v10i3.3035

Moustafa, N., & Slay, J. (2015). UNSW-NB15 : A Comprehensive Data Set for Network Intrusion Detection Systems (UNSW-NB15 Network Data Set). *School Of Engineering And Information Technology University Of New South Wales At The Australian Defence Force Academy Canberra, Australia*. https://doi.org/10.1109/milcis.2015.7348942

Moustafa, N., & Slay, J. (2016). The evaluation of Network Anomaly Detection Systems : Statistical analysis of the UNSW-NB15 data set and the comparison with the KDD99 data set. *Information Security Journal A Global Perspective*, *25*(1-3), 18-31. https://doi.org/10.1080/19393555.2015.1125974

Moustafa, N., Slay, J., & Creech, G. (2019). Novel Geometric Area Analysis Technique for Anomaly Detection Using Trapezoidal Area Estimation on Large-Scale Networks. *IEEE Transactions On Big Data*, *5*(4), 481-494. https://doi.org/10.1109/tbdata.2017.2715166

Alsamiri, J., & Alsubhi, K. (2019). Internet of Things Cyber Attacks Detection using Machine Learning. *International Journal Of Advanced Computer Science And Applications*, *10*(12). https://doi.org/10.14569/ijacsa.2019.0101280

Deze, Z., Huang, H., Hou, R., Rho, S., & Chilamkurti, N. (2021). *Big Data Technologies and Applications : 10th EAI International Conference, BDTA 2020, and 13th EAI International Conference on Wireless Internet, WiCON 2020, Virtual Event, December 11, 2020, Proceedings*. Springer.

Choudhary, S., & Kesswani, N. (2020). *Analysis of KDD-Cup'99, NSL-KDD and UNSW-NB15 Datasets using Deep Learning in IoT. Procedia Computer Science, 167*, 1561-1573. https://doi.org/10.1016/j.procs.2020.03.367

Hasan, M. A. M., Nasser, M., Ahmad, S., & Molla, K. I. (2016). *Feature Selection for Intrusion Detection Using Random Forest. Journal Of Information Security*, 07(03), 129-140. https://doi.org/10.4236/jis.2016.73009

On the KDD'99 Dataset : Support Vector Machine Based Intrusion Detection System (IDS) with Different Kernels. (s. d.). *International Journal Of Electronics Communication And Computer Engineering (TM)*, *4*(4). https://ijecce.org/administrator/components/com_jresearch/files/publications/IJECCE_1835_Final1.pdf