République du Sénégal
Ministère de l'Enseignement Supérieur,
de le Recherche et de l'Innovation
(MESRI)

UVS
UNIVERSITE VIRTUELLE DU SENEGAL
foo nekk foofu la

PÔLE SCIENCES, TECHNOLOGIES ET NUMÉRIQUE

MASTER EN CYBERSECURITE

MÉMOIRE DE MASTER



SUJET : Early Warning Détection et Intelligence Artificielle

PRÉSENTÉ PAR :

Daour DIOP

SOUS LA DIRECTION DE :

M. Chérif DIALLO

Remerciements

Je rends grâce à Dieu le Tout Puissant, Maître des cieux et de la terre, le Très Haut, le Très Miséricordieux de m'avoir donné la force et le courage de mener à bien ce travail et je prie sur son prophète Mohamed(PSL).

J'exprime ma sincère gratitude à mon encadreur Dr Cherif Diallo en tant que professeur avec qui j'ai beaucoup appris tout au long de la formation, pour ses conseils précieux, son aide dans le cheminement de cette étude et pour la peine qu'il s'est donné tout au long de ce travail.

Mes remerciements s'adressent aussi à tous mes professeurs de ce Master en Cybersécurité qui nous ont permis d'acquérir le maximum de connaissance tout au long de cette formation de qualité qui m'a permis d'évoluer dans le domaine de l'informatique surtout en la connaissance de l'environnement de la Cybersécurité.

Je remercie également tous mes camarades du Master en Cybersécurité avec qui j'ai eu une très belle collaboration tout au long de la formation.

Dédicaces

Je dédie de travail à :

Mon très cher Père Moussé Dior DIOP pour m'avoir permis de continuer mes études jusque-là, de m'avoir appris certaines valeurs telles que l'amour du travail, l'esprit de combativité et l'esprit du progrès, son soutien.

Ma très chère mère Maïmouna FALL, une mère exemplaire qui s'est toujours sacrifiée pour sa famille, qui m'a toujours soutenu depuis tout petit et qui m'a toujours encouragé dans tous mes projets.

A toute la famille DIOP et FALL de Kelle à Rufisque, à feu mes grands-parents Moustapha DIOP, Daour FALL et Aïda DIOP, qu'Allah veille sur eux.

A mes frères et sœurs Moustapha, Mbar, Khadija, Thioro, Maïmouna et Marieme, merci pour tout le soutien.

A mes amis et proches Chérif, Badou, Dramane, Baba, Papa, Mady, ma chère Maïmouna, Awa, Ngoné, Mayram, Fatima, Diarra, Khary, Fatou ... La liste serait très longue mais je tenais à tous vous remercier pour tout le soutien et les conseils.

Ce travail n'aurait pas pu être finalisé sans la présence de ces personnes dans ma vie.

Que Dieu vous prête une longue vie et une santé de fer Amine.

Sigles et Abréviations

SIEM: *Système de gestion de bases de données.*

IPS: *Intrusion Prevention System.*

IDS*: Intrusion Detection System*

IA: *Intelligence Artificielle*

ML*: Machine Learning.*

NIDS*: Network Intrusion Detection System*

HIPS: *Host-Based Intrusion Detection System*

KIPS*: Kernel Intrusion Prevention Systems*

WIPS*: Wireless Intrusion Prevention Systems*

ML: *Machine Learning*

TCP: *Transmission Control Protocol*

SOC : *Security Operations Center*

SAP : *Système d'Alerte Précoce*

Avant-propos

L'avancée technologique rapide et la complexification croissante des menaces dans le domaine de la cybersécurité ont donné naissance à un impératif critique : la nécessité de détecter et de répondre rapidement aux potentielles vulnérabilités et attaques. Ce mémoire qui rentre dans le de l'obtention du diplôme d'Ingénieur en Cybersécurité pour le Master en Cybersécurité à l'Université Virtuelle du Sénégal explore le mariage essentiel entre les Systèmes d'Alerte Précoce (SAP) et l'Intelligence Artificielle (IA), deux piliers technologiques qui révolutionnent la manière dont nous abordons la sécurité informatique.

Le monde numérique évolue à une vitesse vertigineuse, et avec cette évolution surgissent des défis de sécurité de plus en plus sophistiqués. Les Systèmes d'Alerte Précoce se positionnent comme une réponse proactive à ces défis, visant à identifier, anticiper et contrer les menaces avant qu'elles ne causent des dommages irréversibles. L'intégration de l'Intelligence Artificielle dans ces systèmes offre une dimension nouvelle, permettant une analyse plus fine des comportements, une détection plus précise des anomalies et une réponse plus rapide aux attaques.

L'objectif principal de ce mémoire est d'explorer en profondeur le fonctionnement des Systèmes d'Alerte Précoce, en mettant particulièrement l'accent sur la manière dont l'Intelligence Artificielle enrichit et optimise ces systèmes. Nous examinerons les composants clés des SAP, les technologies sous-jacentes à leur mise en œuvre, et les avantages qu'apporte l'IA dans la détection et la prévention des menaces.

Ce mémoire est structuré en trois chapitres distincts. Le premier chapitre se penche sur la définition des Systèmes d'Alerte Précoce, exposant leur importance cruciale dans le paysage actuel de la cybersécurité. Le deuxième chapitre explore l'IA et son importance dans de domaine de la sécurité. Enfin le troisième chapitre propose des études de cas concrètes le mariage entre les SAP et l'IA, détaillant les algorithmes, les techniques et les cas d'utilisation qui font de cette union une alliance puissante et des applications pratiques, mettant en lumière les succès et les défis de cette convergence.

Liste des Figures

Figure 1: Système de prévention d'intrusion 4
Figure 2: Système de détection d'intrusion 12
Figure 3: Fonctionnement d'un EDR 16
Figure 4: Fonctionnement d'un SIEM 18
Figure 5: Illustration de l'Intelligence Artificielle 23
Figure 6: Déclaration des variables de l'algorithme 45
Figure 7: Résultat du test de l'algorithme 46

Table des matières

Remerciements iii

Dédicaces iv

Sigles et Abréviations v

Avant-propos vi

Liste des Figures viii

Introduction 1

CHAPITRE 1 : Early Warning System (Early Warning Detection) 3

1.1 Définition : 3

1.2 Mode de Fonctionnement d'un Early Warning System (Early Warning Detection) 3

1.2.1 IPS (Intrusion prevention system) 3

1.2.2 IDS (Intrusion Detection System) : 10

1.2.3 EDR (Endpoint Detection and Response) 15

1.2.4 SIEM (Security Information and Event Managment) 18

1.3 Conclusion 22

CHAPITRE 2 : Intelligence Artificielle et Cybersécurité 23

2.1 Définition 23

2.2 Avantages et inconvénients de l'IA dans la Cybersécurité 25

2.2.1 Les défauts dans la configuration humaine 25

2.2.2 Efficacité humaine et tâches récurrentes 25

2.2.3 Fatigue à l'égard des alertes de menaces 26

2.2.4 Délai de réponse en cas de menace 26

2.2.5 Identification et anticipation des nouvelles menaces 27

2.2.6 Capacité du personnel 28

2.2.7 Adaptabilité 28

2.3 Utilisation de l'IA dans la cybersécurité 29

2.3.1 Comment le Machine Learning est-il utilisé en cybersécurité ? 29

2.3.2 Classification des données 30

2.3.3 Mise en cluster des données 31
2.3.4 Actions recommandées 31
2.3.5 Synthèse des possibilités 32
2.3.6 Prévision prédictive 33
2.4 Exemples d'utilisation du Machine Learning dans le domaine de la cybersécurité 34
2.4.1 Classification et conformité de la confidentialité des données 34
2.4.2 Profils de sécurité des comportements utilisateurs 34
2.4.3 Profils de sécurité des performances système 35
2.4.4 Blocage des robots basés sur le comportement 36
2.5 Conclusion : 37
CHAPITRE 3 : Early Warning Detection ET IA 38
3.1 Définition 38
3.2 Algorithme de Machine Learning Utilisé par les EWS 38
3.3 Cas d'utilisation 44
3.4 Conclusion 47
Conclusion 48
Webographie 49

Introduction

Les avancées technologiques faites pour la digitalisation des tâches ont changé le quotidien de l'homme. Faisant la quasi-totalité de nos activités à travers le numérique, une grande quantité de données importantes et confidentielles est générée et transitée par internet qui est devenu un moyen de communication indispensable pour le travail et le gain en productivité et même pour rendre le quotidien plus facile.

Le cyberespace est ainsi devenu le domaine le plus convoité car étant la plus grande bibliothèque de données et ce flux d'informations augmentant de jour en jour est convoité par des personnes mal intentionnées profitant de la moindre faille du réseau pour se servir à notre insu. Les cyberattaques deviennent de plus en plus fréquentes, rapides et complexes et les pertes occasionnées sont extrêmement coûteuses. Selon Cybersecurity Ventures, 71% des organisations dans le monde ont été victimes de ransomware en 2022 et la cybercriminalité aura infligé au monde plus de 8000 Milliards de dollars en 2023[1]. De ce fait, protéger les réseaux qui sont les chemins par lesquelles transitent toutes ces informations devient une nécessité.

Nous savons que les risques augmentent de la même façon que les avancées technologiques et que les vulnérabilités ''zero day'' ne cessent d'exister, alors nous limiter aux contre-mesures des attaques ou à la simple prévention serait un gros risque.

Pour faire face à ce problème, l'utilisation d'un simple système d'alerte précoce encore appelé « Early Warning System » serait d'un grand apport, mais avec la rapidité et la complexité des attaques, l'associer à l'intelligence artificielle avec ses algorithmes d'apprentissage automatique serait d'un grand atout pour la détection rapide des attaques et la prise de décision pour une réponse qui permettrait de réduire, voire éliminer les dommages.

Ainsi, dans le cadre de ce mémoire pour l'obtention du diplôme d'ingénieur en Cybersécurité, nous avons choisi ce sujet qui résume le parcours de notre formation en Sécurité Informatique et notre spécialisation dans le domaine de l'Intelligence Artificielle. Le travail sera structuré en trois chapitres.

Dans le premier intitulé " Early Warning Detection", nous allons parler de la définition d'un système d'alerte précoce et de son mode de fonctionnement de son domaine d'intervention dans la cybersécurité tout en énonçant quelques risques et attaques cyber avec des exemples de systèmes d'alerte précoce.

Le second chapitre sera une présentation de l'Intelligence Artificielle, quelques de ses exploits, avantages et inconvénients dans la Cybersécurité.

---

[1] *Steve Morgan , rédacteur en chef* Northport, NY – 10 décembre 2022, https://cybersecurityventures.com/stats/

Enfin le troisième chapitre fera office de point de convergence entre un système d'alerte précoce et l'Intelligence Artificielle. Nous y traiterons le fonctionnement d'un Early Warning System utilisant la puissance de traitement des algorithmes de l'IA dans la détection et la prise de décision face à toute sorte de menace liée à cybersécurité avec des exemples dont un cas d'utilisation d'un algorithme de ML et la présentation d'un outils de sécurité lors de la soutenance faisant office de cas pratique.

CHAPITRE 1 : *Early Warning System (Early Warning Detection)*

1.1 Définition :

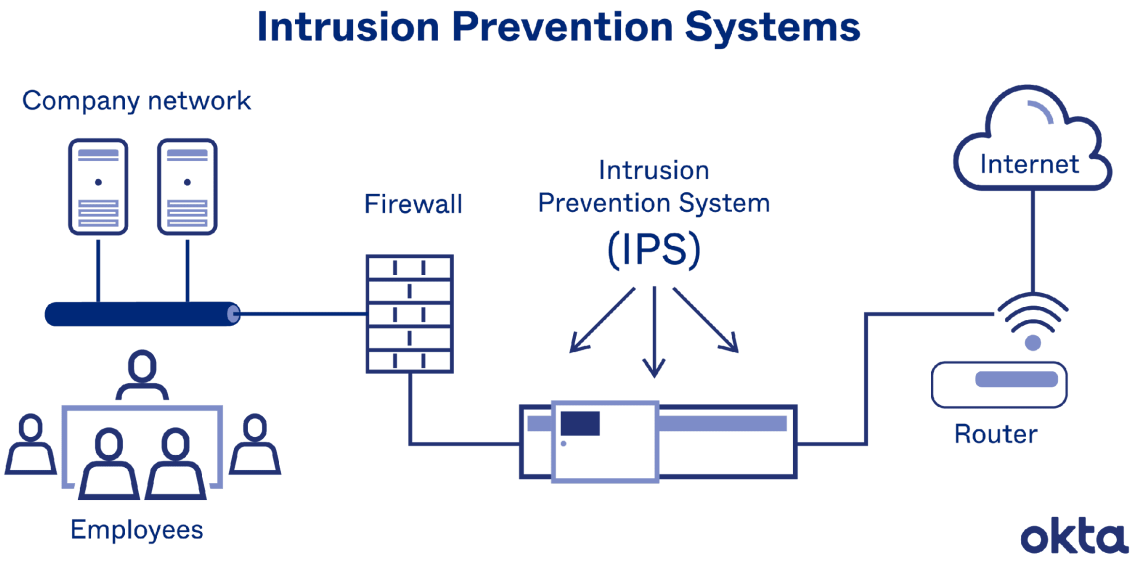
Un Early Warning System ou système d'alerte précoce en cybersécurité désigne un ensemble de procédés et de technologies employés pour repérer et contrer en temps réel d'éventuelles menaces à la sécurité avant qu'elles ne puissent porter atteinte au réseau et aux systèmes d'une entreprise. Cela requiert généralement une surveillance constante des réseaux et des systèmes afin de détecter les signaux révélateurs d'une intrusion et une réaction rapide face aux menaces identifiées dans le but de minimiser les dommages et les perturbations.

Un système d'alerte précoce comprend généralement plusieurs éléments essentiels tels que des systèmes de détection d'intrusion, de systèmes de prévention d'intrusion, des outils de gestion des informations et des événements de sécurité (SIEM), des flux d'informations concernant les menaces ainsi que des plans de réponse aux incidents. Son objectif principal est de fournir aux organisations des informations opportunes et exploitables sur les menaces en émergence, leur permettant ainsi de prendre des mesures proactives pour protéger leurs réseaux et leurs actifs critiques. En installant en place un tel système, les organisations peuvent mieux se prémunir contre les risques potentiels et renforcer leur posture en matière de cybersécurité.

1.2 Mode de Fonctionnement d'un Early Warning System (Early Warning Detection)

1.2.1 IPS (Intrusion prevention system)

Un système de prévention des intrusions (IPS) est une mesure de sécurité réseau qui vise à détecter et prévenir les menaces diffusées. Ces systèmes surveillent en permanence votre réseau, scrutant tout acte potentiellement malveillant et recueillant des informations à ce sujet. En cas de détection d'événements suspects, l'IPS alerte les administrateurs du système et prend des mesures préventives, telles que la fermeture des points d'accès et la reconfiguration des pares-feux, pour empêcher de futures attaques.



<u>Figure 1:</u> Système de prévention d'intrusion

Les solutions IPS peuvent également être utilisées pour identifier les problèmes liés aux politiques de sécurité de l'entreprise, dissuadant ainsi les employés et les invités du réseau de violer les règles incluses dans ces politiques. Étant donné qu'un réseau d'entreprise typique possède de nombreux points d'accès, il est essentiel de disposer d'un moyen de surveiller les signes potentiels d'intrusion, d'incidents et de menaces imminentes. En effet, les menaces réseaux actuelles sont plus en plus sophistiquées et capables d'infiltrer même les solutions de sécurité les plus robustes. Par conséquent, un IPS bien configuré et à jour joue un rôle crucial dans la protection du réseau et des actifs d'une organisation contre les cyberattaques[2].

Un IPS (système de prévention des intrusions) est en effet un IDS (système de détection des intrusions) actif, capable de détecter les activités malveillantes et de prendre des mesures pour bloquer ou prévenir ces attaques.

En détectant des schémas ou des caractéristiques d'activités malveillantes, l'IPS peut réagir rapidement en bloquant les ports, en limitant l'accès à certaines ressources ou en effectuant d'autres actions préventives pour contrecarrer l'attaque.

Il est important de reconnaître que, comme pour les IDS, les IPS ne sont pas infaillibles et peuvent présenter des limitations. Il existe des cas où un IPS peut générer des faux positifs, c'est-à-dire qu'il identifie à tort une activité légitime comme étant malveillante et prend des mesures pour la bloquer. Cela peut entraîner le blocage du trafic légitime, ce qui peut avoir des conséquences négatives sur les opérations normales de l'organisation.

Par conséquent, une configuration appropriée et une surveillance régulière sont essentielles pour minimiser les risques de faux positifs et de faux (où une activité malveillante n'est pas détectée). L'utilisation d'un IPS doit être accompagnée d'une évaluation continue et d'une analyse approfondie des alertes émises, afin de prendre des décisions éclairées sur la

[2] Cyber Edu, Forcepoint https://www.forcepoint.com/fr/cyber-edu/intrusion-prevention-system-ips

meilleure façon de répondre aux menaces tout en évitant des perturbations inutiles pour le trafic légitime.

Les HIPS (Host-Based Intrusion Prevention System) ou les systèmes de prévention des intrusions basés sur l'hôte sont des systèmes de sécurité conçus pour surveiller de manière proactive les postes de travail ou les serveurs, et ils utilisent différentes techniques pour détecter et prévenir les activités malveillantes.

Ils surveillent les processus en cours d'exécution, les pilotes, les bibliothèques de liens dynamiques (.dll), ainsi que d'autres éléments du système d'exploitation et de l'environnement logiciel. En détectant des comportements inhabituels ou des processus suspects qui pourraient indiquer une activité malveillante, le HIPS peut prendre des mesures pour bloquer ou neutraliser immédiatement ces actions dangereuses.

Par exemple, si le HIPS a détecté un processus en train d'effectuer un débordement de tampon (buffer overflow), une technique couramment utilisée dans certaines attaques pour dépasser les limites mémoire d'un programme vulnérable, il peut agir rapidement en stoppant ce processus pour éviter que l'attaque ne réussisse.

En résumé, les HIPS fournissent une couche supplémentaire de défense au niveau de l'hôte, protégeant ainsi contre les attaques ciblant les postes de travail et les serveurs présentant des vulnérabilités logicielles. Grâce à leur capacité à surveiller et à bloquer les comportements suspects, les HIPS contribuent à renforcer la sécurité globale d'un système informatique.

Les NIPS (Network Intrusion Prevention Systems) sont des IPS (Intrusion Prevention Systems) spécialement conçus pour surveiller le trafic réseau en temps réel. Ils analysent le flux de données à la recherche de signes d'activités malveillantes ou d'attaques en cours. Lorsqu'ils détectent une menace, ils peuvent prendre des mesures telles que la terminaison d'une session TCP pour stopper l'attaque et protéger le réseau.

Les NIPS sont un élément essentiel de la sécurité réseau, car ils permettent de détecter et de bloquer rapidement les attaques avant qu'elles ne puissent causer des dommages importants.

Quant aux WIPS (Wireless Intrusion Prevention Systems), ils sont conçus pour la protection des réseaux sans fil. Ils surveillent le trafic des réseaux Wi-Fi pour détecter les activités malveillantes et les tentatives d'intrusion.

En ce qui concerne les KIPS (Kernel Intrusion Prevention Systems), ils sont axés sur la protection du noyau du système d'exploitation. Le noyau étant le cœur du système d'exploitation, les attaques ciblant directement le noyau peuvent être extrêmement dangereuses. Les KIPS sont conçus pour détecter toute tentative d'intrusion ou de compromission du noyau et prendre des mesures pour prévenir de tels événements.

Bien que les NIPS et les WIPS soient largement utilisés dans les environnements de sécurité informatique, les KIPS sont moins courants, car ils développent une expertise plus avancée et peuvent être plus complexes à gérer en raison de leur accès au noyau du système. Cependant, dans certaines situations critiques, les KIPS peuvent fournir une couche supplémentaire de sécurité pour protéger le cœur du système d'exploitation contre les menaces sophistiquées.

Les systèmes de prévention des intrusions fonctionnent en analysant tout le trafic du réseau. Il existe un certain nombre de menaces différentes que les IPS peuvent empêcher, notamment :

- Attaque par déni de service (DDoS)
- Attaque par déni de service distribué
- Divers types de failles de sécurité : Ver, Virus

Lorsqu'un NIPS (Network Intrusion Prevention System) détecte une session TCP utilisée ou une activité malveillante provenant d'une adresse IP source spécifique ou d'un compte utilisateur, il peut prendre des mesures pour empêcher l'accès non éthique à toute application, hôte cible ou ressource réseau. Ces mesures préventives peuvent inclure :

- Terminer la session TCP compromise : Le NIPS peut mettre fin à la session TCP qui a été utilisée ou utilisée de manière malveillante, interrompant ainsi la connexion et interrompant l'attaquant de continuer ses actions nuisibles.

- Bloquer l'adresse IP source : Le NIPS peut bloquer l'adresse IP source de l'attaque, empêcher ainsi toute communication ultérieure provenant de cette adresse. Cela permet de stopper l'attaquant et de le tenir à distance.

- Bloquer le compte utilisateur fautif : Si l'attaque est liée à un compte utilisateur compromis, le NIPS peut bloquer ce compte pour empêcher l'attaquant d'accéder aux ressources avec les identifiants volés.

En prenant ces mesures, le NIPS contribue à renforcer la sécurité du réseau en bloquant les tentatives d'accès non autorisées et en prévenant les activités malveillantes de causer des dommages supplémentaires. Cela permet également de prévenir de futures tentatives d'intrusion de la partie de l'adresse IP ou du compte utilisateur fautif, de renforcer ainsi la défense globale du réseau et des ressources et reprogrammera ou reconfigurera le firewall pour éviter qu'une attaque similaire ne se reproduise à l'avenir.

Un NIPS peut effectivement entreprendre certaines mesures pour entraîner les conséquences d'une attaque malveillante sur le réseau. Voici quelques exemples d'actions qu'un NIPS peut prendre pour supprimer ou remplacer tout le contenu bloqué sur le réseau suite à une attaque :

- Suppression des informations d'en-tête : Si un NIPS a détecté un e-mail erroné qui a réussi à pénétrer dans le réseau, il peut être configuré pour analyser le contenu de l'e-mail et supprimer les informations d'en-tête erronées ou les balises potentielles utilisées pour dissimuler l'origine de l'attaque.

- Retrait des pièces jointes infectées : Si un e-mail ou un fichier des pièces jointes infectées, le NIPS peut être programmé pour identifier ces pièces jointes et les supprimer des serveurs de fichiers ou de courrier électronique contient afin d'empêcher leur diffusion ultérieure.

- Remplacement du contenu malveillant : Dans certains cas, un NIPS peut être configuré pour remplacer le contenu malveillant par un contenu sûr ou un avertissement, de manière à ce que les utilisateurs ne soient pas exposés aux menaces potentielles.

Néanmoins, il est essentiel de noter que la suppression ou le remplacement de contenu défectueux par un NIPS peut présenter des risques et des défis. Si les actions du NIPS ne sont pas correctement configurées, elles pourraient entraîner la perte de données légitimes ou interrompre les services critiques. Par conséquent, la mise en œuvre et la gestion d'un NIPS ont abouti à une approche prudente et à une compréhension approfondie des politiques de sécurité et des besoins spécifiques de l'organisation. Il est également important d'avoir un processus de validation et de vérification pour s'assurer que les actions du NIPS sont correctement effectuées et n'entravent pas les opérations légitimes du réseau.

Un système de prévention des intrusions est généralement configuré pour utiliser un certain nombre d'approches différentes pour protéger le réseau contre les accès non autorisés. On peut citer :

- Approche basée sur les signatures : Cette méthode repose sur la comparaison des signatures connues de menaces spécifiques avec le trafic réseau en temps réel. Une signature est une représentation spécifique d'un modèle ou d'une caractéristique propre à une attaque ou à une activité malveillante particulière. Lorsqu'un système basé sur les signatures détecte une correspondance entre les signatures connues et le trafic en cours d'analyse, il déclenche une alerte pour signaler la présence potentielle d'une menace. Ces signatures peuvent inclure des modèles de virus, des en-têtes de paquets spécifiques, des séquences de commandes utilisées dans les attaques, des comportements de logiciels malveillants, etc.

Les avantages de l'approche basée sur les signatures incluent sa simplicité et sa rapidité de détection pour les menaces déjà connues. Cependant, elle a ses limites, car elle dépend de la disponibilité des signatures à jour pour détecter de nouvelles variantes d'attaques. Les attaques inconnues ou les techniques de contournement peuvent échapper à la détection, car elles ne correspondent pas à des signatures connues.

Pour pallier cette limitation, de nombreux systèmes IDS/IPS utilisent également d'autres techniques, telles que l'analyse comportementale, l'apprentissage automatique et l'analyse heuristique pour détecter des menaces émergentes ou inconnues qui pourraient révéler les signatures traditionnelles. L'approche basée sur les signatures est généralement combinée à ces autres techniques pour fournir une protection plus complète contre les attaques et les activités malveillantes sur le réseau.

- Approche basée sur les anomalies : Contrairement à l'approche basée sur les signatures qui reposent sur la comparaison avec des modèles connus de menaces, l'approche basée sur les anomalies vise à identifier des comportements ou des schémas de trafic qui diffèrent de ce qui est considéré comme normal ou attendu dans le réseau. Ce type d'approche utilise une référence de comportement normal du réseau, généralement basée sur des statistiques et des modèles historiques du trafic, pour établir une "ligne de base" de l'activité typique du réseau. Ensuite, le système surveille en permanence le trafic en cours et compare les schémas actuels aux modèles de référence. Si le système détecte des variations significatives ou des comportements qui s'écartent de la normale, il génère une alerte pour

signaler une activité suspecte ou potentiellement malveillante. Cela peut inclure des connexions provenant d'adresses IP inhabituelles, des photos de trafic inattendus, des tentatives répétées de connexion infructueuses, ou tout autre comportement anormal qui pourrait indiquer une intrusion ou une activité malveillante.

Les avantages de l'approche basée sur les anomalies incluent sa capacité à détecter des attaques inconnues ou émergentes pour lesquelles aucune signature spécifique n'est disponible. Elle peut également être utile pour identifier des activités malveillantes plus des subtiles qui pourraient éviter la détection basée sur les signatures.

Cependant, l'approche basée sur les anomalies peut également générer plus de faux positifs, car certains comportements normaux peuvent être considérés comme anormaux s'ils ne correspondent pas exactement à la ligne de base établie. Pour cette raison, un bon système basé sur les anomalies doit être soigneusement configuré et affiné pour réduire les alertes indésirables tout en identifiant efficacement les véritables menaces potentielles.

- Approche basée sur les politiques Dans cette approche, l'administrateur du système définit des politiques de sécurité qui comportent les composants et les activités autorisés sur le réseau, ainsi que ceux qui sont rejetés comme non autorisés ou suspects. Ces politiques peuvent inclure des règles concernant les types de trafic autorisés, les adresses IP autorisées, les ports autorisés, les protocoles autorisés, etc.

Le système de sécurité examine ensuite le trafic réseau en temps réel et compare les activités aux politiques de sécurité définies. Si une activité enfreint une règle ou correspond à des modèles préoccupants, le système génère une alerte ou prend des mesures préventives pour bloquer l'activité malveillante. L'approche basée sur les politiques offre une grande flexibilité aux administrateurs, car elle leur permet de définir des règles de sécurité spécifiques adaptées aux besoins de leur réseau et de leur organisation. Cela peut inclure des politiques de contrôle d'accès, de prévention d'attaques spécifiques, de blocage de certaines adresses IP, de détection de comportements anormaux, et bien plus encore.

Cependant, cette approche dépend fortement de la qualité et de l'exhaustivité des politiques de sécurité définies par l'administrateur. Des politiques mal configurées ou incomplètes pourraient entraîner des faux positifs ou des faux négatifs, utilisant l'efficacité globale du système de sécurité. Par conséquent, une gestion attentive et une mise à jour régulière des politiques sont essentielles pour maintenir une détection précise et efficace des menaces.

-Inconvénients de L’IPS

L'IPS présente plusieurs inconvénients et limitations qu'il est important de prendre en compte lors de son déploiement et de son utilisation dans un environnement de sécurité réseau. Voici quelques-uns des principaux inconvénients de l'IPS :

- Faux positifs : L'IPS peut générer des alertes pour des activités inoffensives, ce qu'on appelle des faux positifs. Ces alertes erronées peuvent entraîner des perturbations inutiles, une charge de travail supplémentaire pour les équipes de sécurité et une perte de confiance dans l'efficacité du système.

• Faux négatifs : D'un autre côté, l'IPS peut manquer la détection de certaines menaces, ce qu'on appelle des faux négatifs. Les nouvelles variantes d'attaques ou les techniques de contournement perfectionnées peuvent échapper à la détection basée sur les règles ou les signatures utilisées par l'IPS.

• Surcharge du réseau : La surveillance en temps réel et l'analyse du trafic par l'IPS peuvent entraîner une surcharge du réseau, en particulier dans les environnements de trafic réseau intense. Cela peut entraîner une latence ou des ralentissements du réseau.

• Complexité de configuration : La mise en place d'un IPS efficace nécessite une configuration soignée et une compréhension approfondie des politiques de sécurité. La complexité de la configuration peut rendre l'installation et la maintenance de l'IPS plus compliquées.

• Mises à jour effectuées : Pour rester efficace, l'IPS doit être régulièrement mis à jour avec les dernières signatures de menaces et les nouvelles règles de sécurité. Cela nécessite une gestion continue et peut exiger des ressources supplémentaires pour maintenir le système à jour.

• Coûts : Les IPS peuvent être prévus, tant en termes de matériel que de logiciels, ainsi qu'en termes de frais de licence et de maintenance. Les petites organisations avec des budgets limités peuvent trouver difficile d'investir dans un IPS complet.

• Dépendance vis-à-vis des fournisseurs : L'efficacité de l'IPS dépend en grande partie de la qualité des signatures de menaces fournies par les fabricants. Ainsi, l'IPS peut être vulnérable aux nouvelles attaques pour lesquelles il n'existe pas encore de signatures disponibles.

Malgré ces inconvénients, un IPS bien configuré et correctement géré peut apporter une valeur significative en matière de sécurité réseau, en détectant et en prévenant de nombreuses menaces potentielles. Cependant, il est important de prendre en compte ces limitations et de mettre en place des mesures supplémentaires pour renforcer la posture de sécurité globale de l'organisation[3].

1.2.2 IDS (Intrusion Detection System) :

Un système de détection d'intrusion (Intrusion Detection System - IDS) est un outil de sécurité informatique conçu pour surveiller et analyser le trafic réseau et les activités des systèmes informatiques afin de détecter les tentatives d'intrusion, les comportements malveillants ou les activités suspectes. L'objectif principal d'un IDS est de fournir une détection précoce des menaces potentielles et d'alerter les équipes de sécurité pour qu'elles puissent prendre des mesures correctives.

Il existe deux types principaux de systèmes de détection d'intrusion :

[3] Système de prévention d'intrusion — Wikipédia (wikipedia.org)

🕮 L'IDS (Intrusion Detection System) à base de signatures, également connu sous le nom d'IDS basé sur les règles, est une approche de détection d'intrusion qui utilise des bases de données de signatures pour identifier les attaques et les comportements malveillants connus. Ces signatures sont des modèles spécifiques de codes malveillants, de comportements de réseau ou d'activités d'intrusion qui ont été préalablement identifiés et documentés par des experts en sécurité.

Voici comment fonctionne un IDS à base de signatures :

Base de données de signatures : L'IDS dispose d'une base de données de signatures qui contient des milliers de signatures spécifiques, chacune correspondant à un type d'attaque ou de comportement malveillant connu. Ces signatures peuvent inclure des modèles de paquets réseau, des schémas de trafic, des séquences d'instructions ou des valeurs spécifiques dans les journaux d'événements.

Analyse du trafic : L'IDS surveille le trafic réseau en temps réel, analysant les paquets entrants et sortants ainsi que les journaux des systèmes, à la recherche de correspondances avec les signatures de la base de données.

Correspondance des signatures : Lorsqu'un événement ou une activité correspond à une signature de la base de données, l'IDS déclenche une alerte ou une notification pour signaler la détection de l'attaque ou du comportement malveillant.

Génération d'alertes : Une fois qu'une alerte est générée, elle est généralement envoyée à l'équipe de sécurité informatique pour une évaluation plus approfondie et une réponse appropriée.

Les avantages de l'IDS à base de signatures incluent une détection précise des menaces connues pour lesquelles des signatures spécifiques sont disponibles. Il peut rapidement identifier des attaques courantes et aider à prendre des mesures préventives pour bloquer ces menaces connues.

Cependant, un inconvénient majeur de cette approche est qu'elle ne peut détecter que les menaces pour lesquelles des signatures existent. Les nouvelles variantes d'attaques ou les techniques de contournement sophistiquées qui n'ont pas encore été documentées dans la base de données de signatures peuvent échapper à la détection par un IDS à base de signatures.

Pour compléter cette limitation, de nombreuses organisations utilisent également des systèmes de détection d'intrusion basés sur l'analyse comportementale (IDS basé sur les anomalies) et des technologies d'Intelligence Artificielle (IA) pour détecter les menaces inconnues et émergentes qui ne peuvent pas être identifiées par les signatures seules.

🕮 L'IDS (Intrusion Detection System) basé sur l'analyse comportementale, également appelé IDS basé sur les anomalies, est une approche de détection d'intrusion qui se concentre sur la surveillance et l'analyse du comportement du réseau et des systèmes pour détecter les activités anormales ou inhabituelles. Plutôt que de se baser sur des signatures spécifiques d'attaques connues, cette approche vise à identifier les schémas d'activité qui ne correspondent pas aux modèles de comportement normaux ou autorisés.

Voici comment fonctionne un IDS basé sur l'analyse comportementale :

- Établissement d'une ligne de base : Pour commencer, l'IDS observe et enregistre le comportement normal du réseau et des systèmes pendant une période donnée. Cette phase permet de créer une ligne de base des activités normales, qui servira de référence pour identifier les anomalies.

- Surveillance en temps réel : L'IDS surveille en continu les activités du réseau et des systèmes, analysant les journaux d'événements, le trafic réseau et les comportements des utilisateurs.

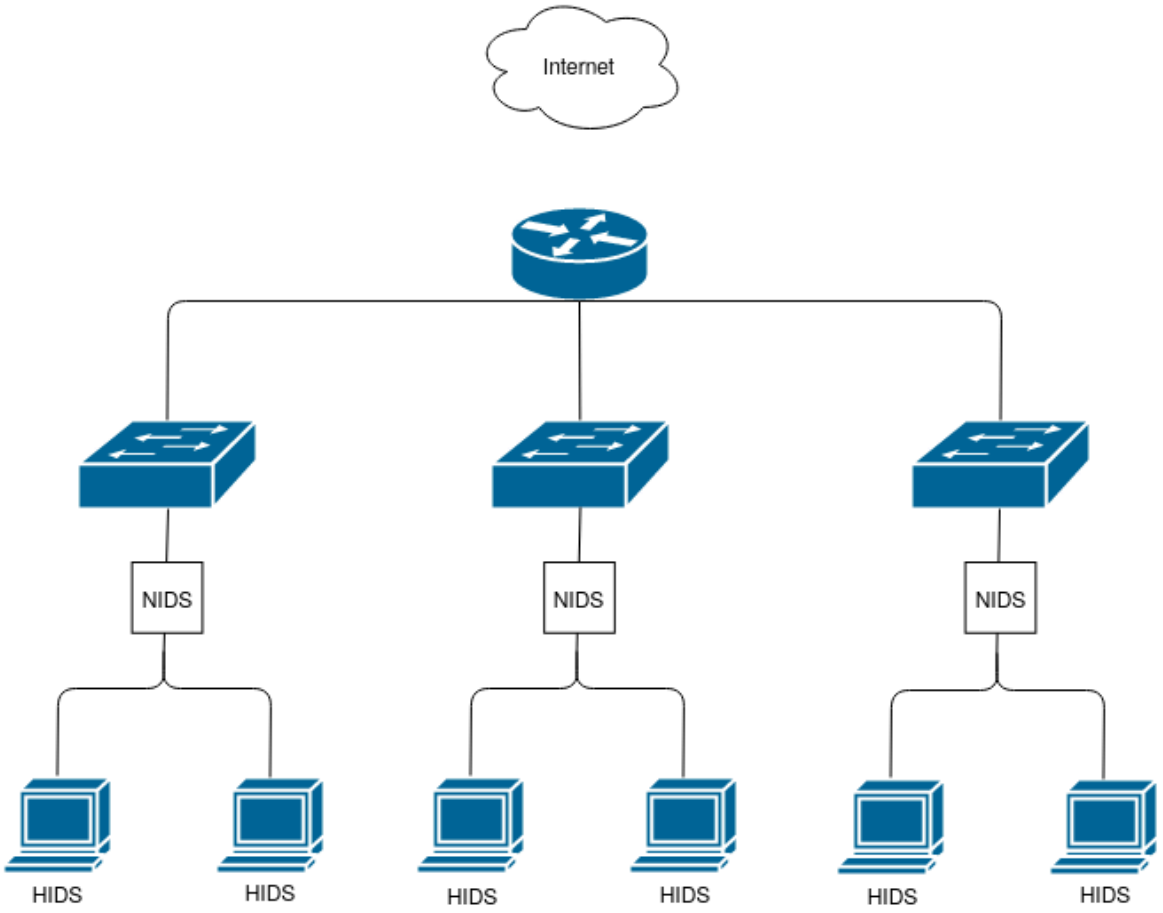
- Détection des anomalies : L'IDS compare les activités en temps réel avec la ligne de base établie. Si des activités ou des comportements déviants sont détectés, qui ne correspondent pas aux schémas de comportement normaux, l'IDS déclenche une alerte pour signaler une potentielle intrusion ou activité malveillante.

- Génération d'alertes : Les alertes générées par l'IDS basé sur l'analyse comportementale sont envoyées à l'équipe de sécurité pour une évaluation plus approfondie et une réponse appropriée.

Les avantages de l'IDS basé sur les anomalies incluent sa capacité à détecter des menaces inconnues, des activités inhabituelles et des comportements malveillants qui ne peuvent pas être identifiés par des signatures spécifiques. Il est efficace pour repérer les attaques zero-day, les activités de piratage sophistiquées et les comportements d'intrusion non conventionnels.

Sans oublier qu'un inconvénient potentiel de cette approche est qu'elle peut générer plus de faux positifs que l'IDS basé sur les signatures, car certaines activités légitimes peuvent être considérées comme anormales si elles s'écartent suffisamment de la ligne de base. Pour atténuer ce risque, les équipes de sécurité doivent mettre en place des seuils appropriés et ajuster la sensibilité de l'IDS basé sur les anomalies en fonction des spécificités de leur environnement.

Dans l'ensemble, l'IDS basé sur l'analyse comportementale est un complément essentiel à l'IDS basé sur les signatures, car il améliore la capacité de détection des menaces inconnues et contribue à renforcer la posture de sécurité globale d'une organisation.



<u>Figure 2:</u> Système de détection d'intrusion

Les IDS peuvent être déployés à différents niveaux du réseau. Nous pouvons citer :

- Un IDS réseau (NIDS - Network Intrusion Detection System) est un type de système de détection d'intrusion qui est spécifiquement conçu pour surveiller le trafic réseau et détecter les activités suspectes ou malveillantes. Contrairement à l'IDS hôte (HIDS), qui est déployé sur des postes de travail ou des serveurs individuels, l'IDS réseau est positionné au niveau du réseau et examine le trafic qui circule entre les différents appareils et systèmes au sein de l'infrastructure. Un IDS réseau (NIDS) fonctionne comme tel:

🕮 Surveillance du trafic : L'IDS réseau surveille en temps réel le trafic réseau à l'aide de sondes ou de capteurs placés à des points stratégiques du réseau. Ces capteurs analysent les paquets de données qui transitent entre les appareils pour détecter toute activité anormale.

🕮 Analyse des paquets : L'IDS réseau examine les en-têtes et les données des paquets de manière non intrusive, en cherchant des modèles de comportement spécifiques, des signatures de malwares connus ou des indicateurs d'attaques potentielles.

🕮 Détection d'anomalies : L'IDS réseau utilise des algorithmes d'analyse comportementale pour détecter des anomalies dans le trafic, telles que des tentatives

répétées de connexion infructueuses, des changements soudains dans le volume de trafic, des motifs de flux inhabituels, etc.

- Corrélation des événements : L'IDS réseau peut corréler les informations provenant de plusieurs capteurs pour obtenir une vue globale du trafic réseau et identifier les schémas d'activité malveillante qui pourraient ne pas être évidents lorsqu'ils sont observés individuellement.

- Génération d'alertes : Lorsqu'une activité suspecte ou malveillante est détectée, l'IDS réseau génère des alertes ou des notifications qui sont transmises à l'équipe de sécurité pour enquête et action.

Les avantages de l'IDS réseau (NIDS) sont qu'il offre une vue globale du trafic réseau, ce qui permet de détecter les menaces qui pourraient passer inaperçues au niveau des systèmes individuels. Il est également capable de détecter les menaces en temps réel, ce qui permet une réponse rapide aux incidents de sécurité.

Cependant, l'IDS réseau peut être limité par sa dépendance vis-à-vis du trafic réseau. S'il n'a pas accès à certaines parties du réseau, il pourrait manquer des activités malveillantes qui se produisent à ces endroits. De plus, l'IDS réseau peut générer des faux positifs en détectant des activités qui ne sont pas nécessairement malveillantes, mais qui peuvent sembler anormales en fonction du contexte spécifique du réseau.

- Un IDS hôte (HIDS - Host-based Intrusion Detection System) qui est un type de système de détection d'intrusion qui est spécifiquement conçu pour surveiller et analyser les activités d'un hôte ou d'un serveur individuel. Contrairement à l'IDS réseau (NIDS) qui se concentre sur le trafic réseau, l'IDS hôte est déployé directement sur chaque poste de travail ou serveur pour examiner les journaux d'événements, les fichiers système et les activités des processus. Un IDS hôte (HIDS) nous permet d'assurer :

- Surveillance des activités du système : L'IDS hôte surveille en temps réel les activités du système d'exploitation, y compris les journaux d'événements, les changements de fichiers, les tentatives de connexion, les activités des utilisateurs, etc.

- Analyse du comportement : L'IDS hôte analyse le comportement des processus et des applications pour détecter les activités malveillantes ou suspectes, telles que la modification de fichiers critiques du système, la création de nouveaux comptes d'utilisateurs inconnus ou l'exécution de programmes potentiellement dangereux.

- Surveillance des fichiers système : L'IDS hôte vérifie l'intégrité des fichiers système importants en utilisant des empreintes numériques ou des mécanismes de détection des changements indésirables. Cela permet de repérer les modifications non autorisées apportées aux fichiers système par des malwares ou des intrus.

- Détection basée sur les signatures : Comme l'IDS réseau, l'IDS hôte peut également utiliser des bases de données de signatures pour détecter des malwares ou des comportements d'intrusion connus.

- Génération d'alertes : Lorsqu'une activité suspecte ou malveillante est détectée, l'IDS hôte génère des alertes ou des notifications pour informer l'équipe de sécurité.

Les avantages de l'IDS hôte (HIDS) sont qu'il offre une visibilité approfondie au niveau de chaque hôte, ce qui permet de détecter les activités malveillantes spécifiques à un système particulier. Il peut également détecter des menaces internes ou des comportements d'utilisateurs non autorisés.

Il serait prudent de noter qu'un inconvénient potentiel de l'IDS hôte est qu'il peut être limité par son accès uniquement aux activités du système sur lequel il est installé. Cela signifie qu'il ne peut pas détecter les menaces qui transitent exclusivement par le réseau et qui n'affectent pas directement l'hôte surveillé.

En résumé, l'IDS hôte (HIDS) est un outil de sécurité important pour surveiller les activités et les comportements au niveau de chaque hôte ou serveur, offrant une couche supplémentaire de détection des menaces et contribuant à renforcer la sécurité des systèmes informatiques au niveau individuel. Il est souvent utilisé en combinaison avec l'IDS réseau (NIDS) pour fournir une approche globale de la détection d'intrusion.

• Un IDS hybride (HIDS/NIDS) qui est une approche de détection d'intrusion qui combine les fonctionnalités des IDS hôte (HIDS - Host-based Intrusion Detection System) et des IDS réseau (NIDS - Network Intrusion Detection System) pour fournir une protection plus complète contre les menaces de sécurité. Cette approche vise à tirer parti des avantages de chaque type d'IDS pour obtenir une vue d'ensemble plus globale et approfondie de la sécurité du réseau et des hôtes.

Avec un IDS hybride (HIDS/NIDS) nous pouvons :

- Surveiller le trafic réseau : Comme un IDS réseau (NIDS), l'IDS hybride surveille le trafic réseau en temps réel à l'aide de capteurs ou de sondes positionnés sur des points stratégiques du réseau. Il analyse les paquets de données et cherche des comportements malveillants connus ou des indicateurs d'attaques.

- Surveiller les activités des hôtes : Comme un IDS hôte (HIDS), l'IDS hybride est déployé directement sur chaque hôte ou serveur pour surveiller les activités du système d'exploitation, les journaux d'événements, les fichiers système et les comportements des processus.

- Corrélation des événements : L'un des aspects clés de l'IDS hybride est sa capacité à corréler les informations provenant des deux sources, à savoir le trafic réseau et les activités des hôtes. En combinant les données des IDS réseau et hôte, l'IDS hybride peut détecter des schémas d'activité qui pourraient ne pas être apparents lorsqu'ils sont observés séparément.

- Génération d'alertes : Lorsque des activités suspectes ou malveillantes sont détectées soit au niveau du réseau, soit au niveau des hôtes, l'IDS hybride génère des alertes pour signaler les menaces potentielles aux équipes de sécurité.

Les avantages de l'IDS hybride (HIDS/NIDS) résident dans sa capacité à fournir une détection plus complète et précise des menaces de sécurité. En combinant les informations

du trafic réseau et des activités des hôtes, il peut identifier les comportements d'attaques plus complexes qui pourraient passer inaperçus par les IDS séparés.

Cependant, l'IDS hybride peut également être plus complexe à déployer et à gérer en raison de la nécessité de coordonner les données provenant de sources différentes. En outre, cela peut entraîner une augmentation du volume d'alertes, ce qui nécessite une bonne gestion des alertes pour éviter les faux positifs et garantir une réponse appropriée aux véritables incidents de sécurité.

Les IDS sont un élément essentiel de la stratégie de sécurité car ils contribuent à renforcer la détection des menaces et à améliorer la réponse aux incidents de sécurité, permettant ainsi de protéger les systèmes et les données contre les attaques potentielles[4].

1.2.3 EDR (Endpoint Detection and Response)

Les antivirus traditionnels ont joué un rôle essentiel dans la sécurité informatique en détectant et en neutralisant de nombreuses menaces connues. Avec l'évolution constante des attaques et l'émergence de menaces plus simplifiées et ciblées, il est devenu évident qu'une approche basée uniquement sur des signatures et une analyse individuelle poste par poste peut présenter des limitations.

L'une de ces limitations est effectivement la corrélation des événements et des activités à travers l'ensemble du parc informatique. Dans un environnement de sécurité actuel, les attaques peuvent se propager rapidement et discrètement, passant d'un poste à un autre ou exploitant des vulnérabilités spécifiques dans le réseau. Il est donc essentiel d'avoir une visibilité globale sur l'ensemble du parc pour comprendre la trajectoire de l'infection initiale, ainsi que les mouvements latéraux des attaquants.

Cependant, cette corrélation peut être complexe et chronophage, surtout si les outils de sécurité et les solutions de gestion de la sécurité ne sont pas intégrés ou ne se séparent pas facilement des informations entre eux. L'absence de visibilité globale peut entraîner des retards dans la détection et la réponse aux incidents, ce qui permet aux agresseurs de rester actifs plus longtemps et d'augmenter les risques pour l'organisation. Pour faire face à ces limitations, les entreprises adoptent des solutions de sécurité avancées telles que les solutions EDR (Endpoint Detection and Response) qui fournissent une visibilité approfondie au niveau de chaque poste de travail. Ces solutions permettent une détection plus précise des activités suspectes et aident à comprendre l'étendue d'une attaque sur l'ensemble du réseau.

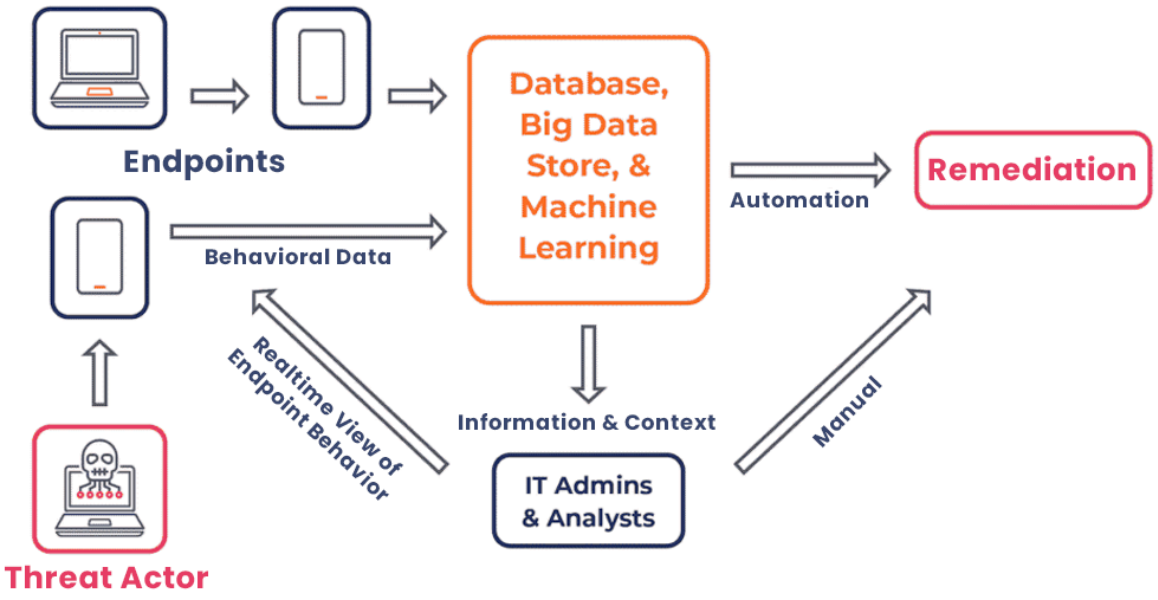
L'EDR est en effet un outil de sécurité complémentaire de l'antivirus de nouvelle génération (Next-Generation Antivirus - NGAV). Alors que les solutions antivirus traditionnelles se concentrent principalement sur la détection des malwares connus en se basant sur des signatures, l'EDR va au-delà en fournissant une visibilité plus approfondie au niveau de chaque poste de travail ou terminal dans un réseau.

Contrairement aux solutions de sécurité au niveau du réseau, qui surveillent et analysent le trafic à un niveau plus global, l'EDR est déployé directement sur chaque terminal individuel.

[4] https://fr.wikipedia.org/wiki/Syst%C3%A8me_de_d%C3%A9tection_d%27intrusion

Cela permet à l'EDR de collecter et d'analyser en temps réel les données sur les activités et les comportements des processus, des fichiers et des utilisateurs sur chaque poste de travail.

Comment fonctionne l'EDR ?



<u>Figure 3</u>: Fonctionnement d'un EDR

En utilisant des techniques telles que l'analyse comportementale, l'apprentissage machine et la détection heuristique, l'EDR est capable de détecter les activités suspectes et les comportements malveillants, y compris les menaces inconnues ou zero-days. Cela signifie qu'il peut détecter des attaques basées sur de nouvelles vulnérabilités ou des méthodes de contournement qui n'ont pas encore de signatures connues.

Un EDR est caractérisé par ses capacités de détection, d'investigation et enfin de remédiation.

a) Detection :

L'EDR est capable de surveiller l'exploitation de failles de sécurité en surveillant le noyau et les différents services habituellement ciblés notamment chez Windows. Cette capacité de surveillance et la corrélation d'évènements lui permettent de reconnaître des méthodes et habitudes qu'ont les hackers et dont il est plus difficile de se prémunir.

L'analyse comportementale est un autre point étudié par les EDR et qui permet de reconnaître des comportements déviant d'une norme, après une phase d'apprentissage. Grâce à cette analyse, l'EDR peut émettre des alertes qui seront vérifiées et renforceront l'apprentissage. L'intérêt de cette technique est qu'elle permet de stopper un attaquant dans son élan : si un fichier PDF contient un script qui ouvre powershell et ouvre une connexion sur un port classique d'un serveur extérieur au SI alors cette suite d'action sera considérée comme anormale et va être bloquée par l'EDR. Cette visibilité est une grande force de l'EDR car elle permet une remédiation à la source de l'infection.

b) Investigation :

Comme mentionné plus haut, l'EDR permet d'observer des suites d'actions dont le résultat est des plus douteux. Cette visibilité sur les processus est d'une très grande aide pour faire de l'investigation : les actions sont corrélées et remontées dans une plateforme centralisée qui permet d'étendre l'apprentissage observé d'un poste à tous les autres. Ainsi, si une attaque est détectée sur cinq terminaux, alors la console centralisée va faire redescendre l'information sur tous les autres.

Grâce à cette plateforme, un Security Operations Center (SOC) est capable de savoir immédiatement le nombre de postes sont touchés et d'assurer la traçabilité. C'est un formidable outil d'investigation qui s'interfacera à un Security Information and Event Management (SIEM) en offrant de la visibilité sur les terminaux. D'autant plus qu'un SOC pourra se servir de l'EDR pour récupérer toutes les informations en cas d'attaque à distance.

c) Remédiation:

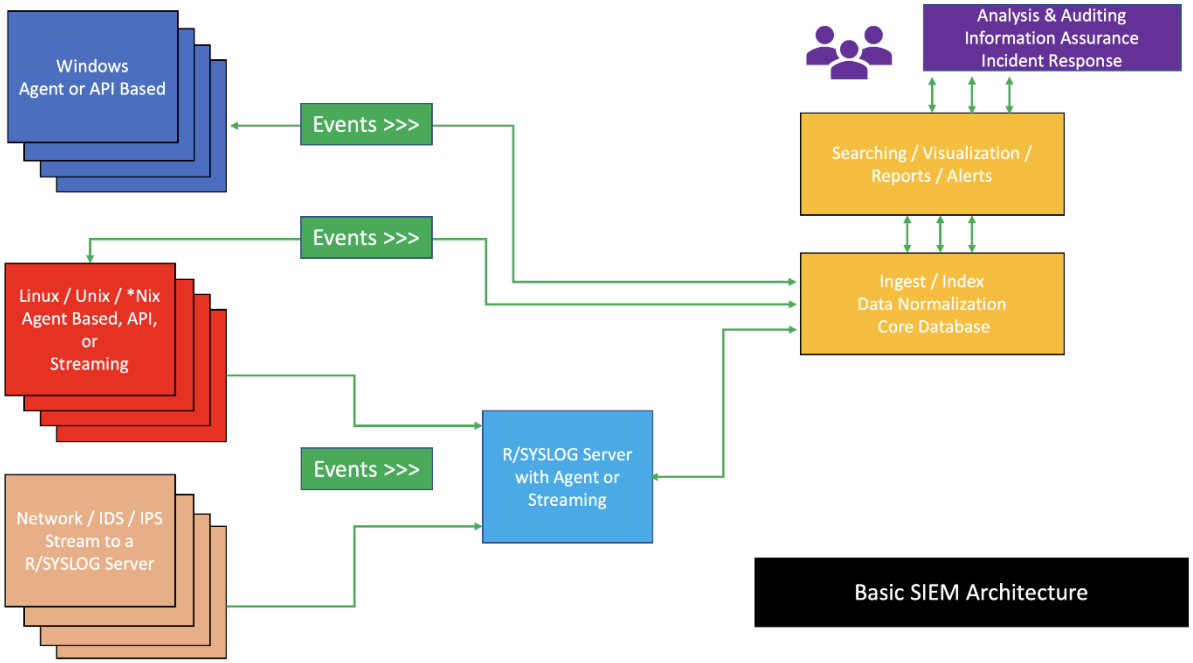
En ce qui concerne la remédiation, l'EDR a des capacités similaires à celles d'un antivirus de nouvelle génération et peut notamment bloquer, supprimer et mettre en quarantaine des fichiers. Les équipes de sécurité pourront aussi s'appuyer sur l'outil pour faire du nettoyage en profondeur afin de corriger les mauvaises actions faites par un malware. De plus, certains EDR permettent aux analystes du SOC de faire un contrôle à distance sur un terminal qui nécessiterait une investigation plus poussée encore.

L'EDR n'est pas une solution autonome. C'est un excellent complément qui s'intégrera parfaitement avec un antivirus classique, un SIEM et de la sécurité du réseau. Il permet d'étendre la visibilité du SI jusqu'au terminaux et ainsi d'améliorer et d'étendre la sécurité du SI[5].

1.2.4 SIEM (Security Information and Event Managment)

Les SIEM (Security Information and Event Management) ou système de gestion des événements et des informations de sécurité sont des outils essentiels dans la stratégie de sécurité informatique d'une organisation, car ils jouent un rôle clé dans la détection précoce des menaces et dans la réponse aux incidents de sécurité. Pour se faire, le SIEM tire ses informations à partir diverses sources qui peuvent inclure des systèmes d'exploitation, de pare-feu, des antivirus, des dispositifs de sécurité réseau, des applications, des serveurs, des objets connectés et bien d'autres.

[5] Orange Cyberdefence https://www.orangecyberdefense.com/fr/solutions/protection-des-mobiles-et-des-endpoints/endpoint-detection-and-response-pourquoi-ledr



<u>Figure 4:</u> Fonctionnement d'un SIEM

Les SIEM ont pour principal objectif de fournir une visibilité globale sur l'environnement de sécurité de l'organisation et de détecter toutes activités suspectes ou malveillantes. Les SIEM collectent, analysent et corrèlent en temps réel les journaux et les événements de sécurité provenant de diverses sources au sein du réseau et des systèmes de l'organisation. Parmi les rôles d'un SIEM nous pouvons citer :

- Collecte de données : La collecte de données est une étape essentielle du fonctionnement d'un SIEM (Security Information and Event Management). Les SIEM collectent une variété de journaux et d'événements de sécurité provenant de différentes sources au sein du réseau et des systèmes de l'organisation. Les principales sources de données collectées par un SIEM sont :

🕮 Journaux des systèmes d'exploitation : Les SIEM collectent les journaux d'événements des systèmes d'exploitation des serveurs, postes de travail et appareils réseau. Ces journaux peuvent contenir des informations sur les connexions, les erreurs, les activités des utilisateurs, les changements de configuration et bien plus encore.

🕮 Logs des applications : Les SIEM collectent également les journaux générés par les applications et les services déployés dans l'environnement de l'organisation. Ces journaux peuvent inclure des informations sur les activités des applications, les erreurs, les tentatives d'accès, etc.

🕮 Logs des dispositifs de sécurité : Les SIEM collectent les journaux des pare-feu, des systèmes de prévention d'intrusion (IPS), des systèmes de détection d'intrusion (IDS), des antivirus, des solutions de filtrage de contenu, etc. Ces logs fournissent des informations sur les tentatives d'intrusion, les événements de sécurité, les tentatives de blocage de trafic malveillant, etc.

- Logs réseau : Les SIEM collectent les journaux de trafic réseau, qui incluent des informations sur les communications entre les appareils, les flux de données, les adresses IP source et destination, les ports utilisés, etc.

- Journaux d'authentification : Les SIEM collectent les journaux d'authentification et d'accès, tels que les tentatives de connexion réussies ou échouées, les changements de mot de passe, les comptes utilisateurs créés ou désactivés, etc.

- Logs d'audit : Certains systèmes et bases de données génèrent des logs d'audit qui enregistrent les modifications apportées aux données sensibles, telles que les enregistrements d'accès aux bases de données.

Une fois que les données sont collectées, le SIEM les stocke dans une base de données sécurisée pour une analyse ultérieure. L'analyse en temps réel des données collectées permet de détecter les activités malveillantes ou suspectes, d'identifier les schémas d'attaques et de générer des alertes pour signaler les menaces potentielles.

- Analyse en temps réel :

L'analyse en temps réel est l'une des fonctionnalités clés d'un SIEM. Cette analyse permet au SIEM de surveiller continuellement les données collectées en temps réel, d'identifier rapidement les activités suspectes ou malveillantes et de générer des alertes pour signaler les menaces potentielles. L'analyse en temps réel du SIEM fonctionne comme tel:

- Collecte de données en temps réel : Le SIEM collecte en continu les journaux et les événements de sécurité provenant de diverses sources au sein du réseau et des systèmes de l'organisation. Ces données sont transmises en temps réel aux serveurs du SIEM pour analyse.

- Prétraitement des données : Avant l'analyse, les données collectées peuvent être prétraitées pour nettoyer, normaliser et enrichir les informations. Cela permet d'assurer que les données sont cohérentes et prêtes pour l'analyse.

- Analyse des journaux et des événements : Le SIEM analyse les journaux et les événements en temps réel pour détecter les activités suspectes ou malveillantes. L'analyse peut inclure l'utilisation de règles et de filtres préconfigurés pour identifier les comportements indésirables connus.

- Suivi en temps réel : En plus de la détection des menaces, le SIEM peut permettre un suivi en temps réel de l'état de la sécurité du réseau et des systèmes à l'aide de tableaux de bord et de visualisations. Cela permet aux équipes de sécurité de surveiller l'activité et de réagir rapidement aux incidents.

- Corrélation des événements : La corrélation des événements est une fonctionnalité importante d'un SIEM qui permet de détecter les schémas d'activité malveillante en reliant des événements apparemment indépendants provenant de différentes sources au sein du réseau et des systèmes de l'organisation. Cette corrélation permet de découvrir des menaces complexes et de fournir une vue globale des incidents de sécurité. La corrélation des événements du SIEM se fait par :

- Collecte de données : Le SIEM collecte en temps réel les journaux et les événements de sécurité provenant de diverses sources, telles que les systèmes d'exploitation, les applications, les dispositifs de sécurité réseau, les pare-feu, les systèmes de détection d'intrusion (IDS), les systèmes de prévention d'intrusion (IPS), etc.

- Prétraitement des données : Avant la corrélation, les données collectées peuvent être prétraitées pour normaliser, nettoyer et enrichir les informations. Cela garantit que les données sont cohérentes et prêtes à être analysées.

- Identification des événements liés : Le SIEM analyse les événements en temps réel pour identifier les activités potentiellement liées. Par exemple, il peut détecter des tentatives de connexion échouées sur un pare-feu suivies de tentatives de connexion réussies sur un serveur, ce qui pourrait indiquer une tentative d'intrusion.

- Analyse des séquences d'événements : Le SIEM analyse également les séquences d'événements dans le temps pour détecter des schémas d'activité malveillante. Par exemple, il peut repérer une série d'actions inhabituelles, telles que l'accès à des fichiers sensibles suivi d'une exfiltration de données.

- Détection de compromission : En reliant des événements apparemment indépendants, le SIEM peut identifier des indicateurs de compromission (IoC - Indicators of Compromise) et des modèles d'attaques sophistiquées qui pourraient passer inaperçus par des analyses individuelles.

- Génération d'alertes :

La génération d'alertes est l'un des atouts essentiels d'un SIEM. Lorsque le SIEM analyse en temps réel les données collectées et détecte des activités suspectes ou malveillantes, il génère des alertes ou des notifications pour informer les équipes de sécurité de la présence de menaces potentielles. Cette génération se fait en :

- Définition de seuils : Le SIEM peut être configuré pour définir des seuils spécifiques pour les événements. Par exemple, il peut générer une alerte lorsqu'un grand nombre de tentatives de connexion échouées sont détectées dans un court laps de temps.

- Classification des alertes : Les alertes générées par le SIEM peuvent être classées en fonction de leur gravité ou de leur niveau de priorité. Cela aide les équipes de sécurité à hiérarchiser leurs réponses en fonction de l'importance des incidents.

- Envoi des alertes : Une fois qu'une activité suspecte est détectée, le SIEM génère une alerte et envoie une notification aux équipes de sécurité concernées. Cela peut se faire par le biais de messages d'alerte dans l'interface du SIEM, d'e-mails, de SMS ou d'autres moyens de communication.

- Réponse aux alertes : Les équipes de sécurité peuvent ensuite prendre des mesures pour enquêter sur l'alerte, confirmer si elle représente une véritable menace, atténuer les risques et remédier à l'incident de sécurité.

• Gestion des événements de sécurité (rapports et tableau de bord):

La gestion des événements de sécurité dans un SIEM (Security Information and Event Management) comprend la création de rapports et de tableaux de bord pour visualiser, analyser et présenter les informations liées à la sécurité de l'organisation. Ces rapports et tableaux de bord offrent une vue globale et claire de l'état de la sécurité, des incidents détectés, des tendances de sécurité, des performances du SIEM, etc. La gestion des événements de sécurité se fait par :

- Collecte de données : Le SIEM collecte en temps réel les journaux et les événements de sécurité provenant de diverses sources au sein du réseau et des systèmes de l'organisation.

- Stockage des données : Les données collectées sont stockées dans une base de données sécurisée pour une analyse ultérieure et une rétention à long terme.

- Analyse des données : Le SIEM analyse les données collectées pour détecter les activités suspectes, corréler les événements, identifier les indicateurs de compromission et les tendances de sécurité.

- Création de rapports : Le SIEM génère des rapports détaillés sur différents aspects de la sécurité. Ces rapports peuvent inclure des informations sur les incidents de sécurité, les activités d'authentification, les tentatives d'intrusion, les comportements anormaux, les performances du SIEM, etc. Les rapports peuvent être configurés pour être générés à intervalles réguliers (quotidiennement, hebdomadairement, mensuellement, etc.) ou à la demande.

- Tableaux de bord : Le SIEM crée également des tableaux de bord interactifs qui présentent visuellement les données de sécurité sous forme de graphiques, de diagrammes, de jauges et d'autres éléments. Ces tableaux de bord offrent une vue d'ensemble de la sécurité en temps réel, facilitant la détection rapide des problèmes et l'analyse des tendances.

- Personnalisation : Les rapports et les tableaux de bord du SIEM peuvent être personnalisés pour répondre aux besoins spécifiques de l'organisation. Les équipes de sécurité peuvent sélectionner les types d'informations à inclure dans les rapports, choisir les paramètres de corrélation et de filtrage, et personnaliser l'apparence des tableaux de bord.

🕮 Diffusion des rapports : Les rapports peuvent être diffusés aux parties prenantes pertinentes au sein de l'organisation, telles que les équipes de sécurité, les responsables de la conformité, la direction et les équipes informatiques. Cela permet de partager les informations sur la sécurité et d'améliorer la transparence.

La gestion des événements de sécurité avec les rapports et tableaux de bord du SIEM est essentielle pour surveiller l'état de la sécurité, analyser les incidents, détecter les menaces potentielles, répondre rapidement aux incidents de sécurité et améliorer la posture globale de sécurité informatique de l'organisation. Ces rapports et tableaux de bord aident à prendre des décisions éclairées en matière de sécurité et à renforcer la résilience face aux menaces de plus en plus sophistiquées.

- Intégration de Threat Intelligence :

Certains SIEM intègrent des renseignements sur les menaces provenant de sources externes, telles que les centres de veille et de renseignement sur les menaces, pour enrichir la détection des menaces et identifier de nouvelles formes d'attaques.

En combinant toutes ces techniques, les SIEM permettent aux organisations de détecter les menaces de sécurité plus rapidement, de répondre aux incidents plus efficacement et de renforcer leur posture de sécurité globale. La détection précoce des menaces assurée grâce à un SIEM est essentielle pour prévenir les incidents de sécurité, limiter les dommages potentiels et la résilience de l'organisation face aux cyberattaques[6].

1.3 Conclusion

Un système d'alerte précoce est un outil révolutionnaire pour la cybersécurité. Il est d'une très grande utilité pour les organisations et son automatisme dans la prévention, la detection et la réponse aux menaces serait plus performante lorsqu'elle sera combinée avec l'intelligence artificielle qui, de par sa puissance inestimable de calcul pourrait faciliter tous les processus.

[6]https://www.microsoft.com/fr/security/business/security-101/what-is-siem#:~:text=Un%20syst%C3%A8me%20de%20gestion%20des,elles%20ne%20perturbent%20leurs%20activit%C3%A9s.

2.1 Définition

L'Intelligence Artificielle peut être définie comme la reproduction de l'intelligence humaine. Cela va du programme d'analyse médicale détectant une maladie jusqu'à celui capable de conduire une voiture en toute autonomie. Parmi ces applications, nous pouvons distinguer celles dont les règles ont été fixées à l'avance par des experts pour une réponse spontanée à une demande de celles dotées d'une faculté à adapter leur comportement en fonction de la situation. Pour ce second cas, les termes d'apprentissage automatique, ou de Machine Learning (ML) sont utilisés. Ces systèmes se basent sur un grand nombre de données, manipulées à l'aide d'ordinateurs de plus en plus puissants, qu'ils analysent pour reconnaître automatiquement des motifs servant de base à des décisions.

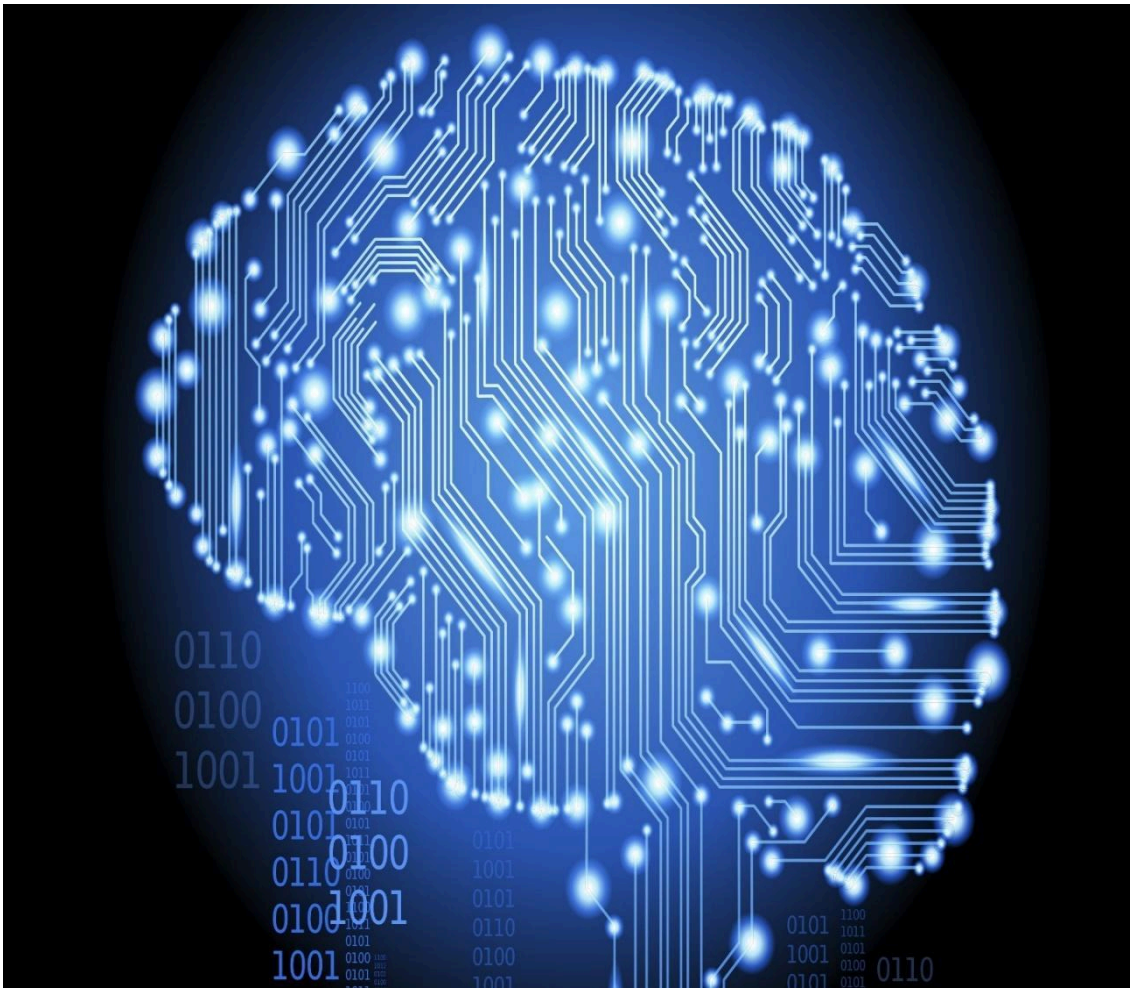


Figure 5: Illustration de l'Intelligence Artificielle

Il existe différentes méthodes d'apprentissage, non exhaustives : apprentissage supervisé, non-supervisé ou apprentissage par renforcement.

Avec leur mécanisme d'apprentissage et leur faculté à adapter leurs seuils de décision de manière évolutive, ces systèmes ont une différence révolutionnaire avec les systèmes historiques. Aussi, il y'a une certaine nuance sur la compréhension entre Intelligence Artificielle et Machine Learning, y compris pour évoquer le sujet de la cybersécurité. Ce principe ne fait pas exception à la règle et, derrière les mots « Intelligence Artificielle », il se focalise plus particulièrement sur la gestion des risques de cybersécurité liés à l'utilisation du Machine Learning.

La cybersécurité basée sur l'IA, avec l'aide du Machine Learning, est conçue pour être un outil puissant dans un avenir proche. Comme dans les autres secteurs, l'intervention humaine est a toujours été essentielle et irremplaçable en matière de sécurité. Même si la cybersécurité s'appuie fortement sur l'interaction humaine, la technologie devient exponentiellement de plus en plus efficace sur des tâches spécifiques.

Les avancées technologiques nous mènent doucement vers une complémentarité avec les interventions humaines en les rendant plus efficaces. Certains axes de recherche s'inscrivent au cœur de ces avancées :

- L'intelligence artificielle (IA) : qui est un domaine de l'informatique qui vise à développer des systèmes informatiques capables de simuler des processus cognitifs humains tels que l'apprentissage, la prise de décision, la reconnaissance de la parole, la vision, la compréhension du langage, et bien d'autres. L'IA cherche à donner aux machines la capacité d'apprendre à partir de données, de s'adapter aux nouvelles informations et de prendre des décisions en se basant sur ces connaissances, tout comme le ferait un esprit humain mais plus rapidement.

- Le Machine Learning (ML) : C'est une approche de l'IA qui permet aux ordinateurs d'apprendre à partir de données sans être explicitement programmés. Les algorithmes de Machine Learning identifient des modèles dans les données et utilisent ces modèles pour prendre des décisions, faire des prédictions ou résoudre des problèmes. L'intervention humaine est toujours requise pour le suivi et la programmation des modifications et mise à jour. Le Machine Learning est probablement la discipline de l'IA la plus pertinente à ce jour utilisée pour la Cybersécurité.

- Le Deep Learning (DL) : Il fonctionne de la même manière que le ML. C'est une sous-catégorie du Machine Learning qui s'inspire du fonctionnement des réseaux de neurones du cerveau humain. Les réseaux de neurones profonds sont utilisés pour résoudre des problèmes complexes tels que la reconnaissance d'images, le traitement du langage naturel et la compréhension du langage. Il prend des décisions à partir de schémas déjà vus, mais effectue des ajustements de manière autonome. Dans le domaine de la cybersécurité, le Deep Learning relève du Machine Learning, sur lequel nous ferons focus davantage.

2.2 Avantages et inconvénients de l'IA dans la Cybersécurité

L'IA et la cybersécurité ont été présentées comme révolutionnaires et bien plus liées que l'on ne pensait. Néanmoins, ce côté révolutionnaire n'est que partiel et doit être considéré avec

mesure. En réalité, les améliorations progressives sont d'une grande utilité mais peuvent aussi causer des dommages de par une mauvaise utilisation. Pour mettre les choses en perspective, ce qui peut sembler progressif en comparaison avec un avenir tout à fait autonome, va aujourd'hui au-delà de ce dont nous avons été capables par le passé.

De ce fait, dans l'étude et la recherche des implications possibles avec la sécurité dans le ML et l'IA, il est important de cadrer les difficultés actuelles inhérentes à la cybersécurité. Il existe de nombreux processus et aspects que nous avons longtemps considérés comme normaux, qui peuvent être traités sous l'égide des technologies IA.

2.2.1 Les défauts dans la configuration humaine

L'erreur humaine l'une des plus grandes faiblesses de la cybersécurité. Par exemple, une configuration système appropriée peut être incroyablement difficile à gérer, même avec d'importantes équipes informatiques impliquées. Dans la course à l'innovation constante, la sécurité informatique est devenue plus hiérarchisée que jamais. Des outils adaptés peuvent aider les équipes à identifier et à atténuer les problèmes survenant lorsque les systèmes réseau sont remplacés, modifiés et mis à jour.

Lorsque nous observons la manière dont les dernières infrastructures Internet comme le Cloud Computing peuvent servir de remplaçant aux anciennes structures locales, il devient crucial pour un département informatique de garantir la compatibilité des systèmes d'entreprise pour les protéger. Les processus manuels d'évaluation de la sécurité des configurations épuisent les équipes, qui jonglent avec les mises à jour interminables et les tâches de support quotidiennes ordinaires. De ce fait, grâce à une automatisation intelligente et adaptative, les équipes peuvent bénéficier d'une efficacité importante pour la réponse aux menaces détectées. Elles peuvent prendre conseil sur les options de traitement, voire disposer de systèmes pour régler automatiquement les paramètres si nécessaire.

2.2.2 Efficacité humaine et tâches récurrentes

L'efficacité humaine est un autre défi inhérent au secteur de la cybersécurité. Nul ne peut assurer qu'un processus manuel soit parfaitement répétable à chaque fois, notamment dans des environnements dynamiques comme les nôtres. La configuration individuelle des nombreux terminaux d'une entreprise fait partie des tâches les plus chronophages. Même après la configuration initiale, le département informatique est amené à vérifier à nouveau les mêmes machines ultérieurement pour corriger des configurations inappropriées ou obsolètes ne pouvant pas être rectifiées dans le cadre de mises à jour à distance, ce qui devient une corvée de plus.

Une observation de plus, lorsque les employés doivent réagir face à des menaces, la portée desdites menaces peut rapidement évoluer. Quand les interventions humaines peuvent être ralenties par des imprévus liés à leur vie, un système basé sur l'IA et le Machine Learning peut fonctionner avec un retard minime.

2.2.3 Fatigue à l'égard des alertes de menaces

La fatigue à l'égard des alertes de menaces constitue une autre faiblesse si elle n'est pas soigneusement gérée. Les surfaces des attaques s'élargissent à mesure que les niveaux de sécurité susmentionnés deviennent de plus en plus sophistiqués et tentaculaires. De nombreux systèmes de sécurité sont configurés pour réagir à une multitude de problèmes

connus, avec une foule d'alertes purement machinales. En conséquence, ces invites individuelles laissent le département informatique analyser les décisions potentielles et prendre les mesures qui s'imposent.

Le volume élevé d'alertes générées par les systèmes de sécurité, y compris les SIEM, les IDS et autres solutions de détection, peut entraîner une surcharge d'informations pour les équipes de cybersécurité. Cette surcharge conduit souvent à la fatigue décisionnelle, où les membres du personnel responsables de la cybersécurité peuvent devenir submergés par la quantité d'alertes à traiter, ce qui peut compromettre leur capacité à prendre des décisions éclairées et rapides et la priorisation des menaces en cas d'incident.

Exploiter l'IA dans le domaine de la cybersécurité peut permettre aux départements informatiques de gérer davantage de menaces de manière efficace et pratique. Un regroupement par étiquetage automatisé peut permettre de simplifier la gestion de chacune de ces menaces. En outre, certains problèmes peuvent être résolus par l'algorithme de Machine Learning lui-même.

2.2.4 Délai de réponse en cas de menace

Le délai de réponse est un facteur critique dans l'efficacité d'une équipe de cybersécurité. Avec l'évolution des techniques d'attaque et l'utilisation croissante de l'automatisation par les cybercriminels, les menaces évoluent rapidement et les attaques peuvent se propager à une vitesse alarmante. Les attaquants cherchent constamment à exploiter les vulnérabilités et à contourner les mesures de sécurité mises en place. Les points importants à considérer sont les suivants :

- Rapidité de propagation des attaques : Les attaques se propagent rapidement, notamment grâce à l'automatisation, ce qui signifie que les équipes de cybersécurité ont un temps limité pour détecter, analyser et répondre à une attaque avant que des dommages importants ne soient causés.

- Automatisation des attaques : L'automatisation est devenue un élément clé dans les cyberattaques, ce qui permet aux attaquants de lancer des campagnes d'attaque à grande échelle et de manière coordonnée, rendant plus difficile pour les équipes de sécurité de suivre le rythme.

- Difficulté de détection : Certaines attaques peuvent être difficiles à détecter, surtout si les attaquants utilisent des techniques furtives pour échapper aux systèmes de détection traditionnels.

- Priorisation des alertes : Face à un grand nombre d'alertes, il est essentiel pour les équipes de cybersécurité de prioriser rapidement les incidents les plus critiques et de se concentrer sur les attaques qui représentent une menace immédiate pour l'organisation.

Pour faire face à ces défis, les organisations peuvent adopter des mesures proactives pour renforcer leur posture de sécurité, telles que :

- Automatisation de la réponse aux incidents : L'utilisation de l'automatisation pour répondre rapidement à certaines attaques peut aider à limiter les dégâts avant qu'ils ne deviennent incontrôlables.

- Threat Hunting : Les équipes de cybersécurité peuvent mener des activités de "Threat Hunting" pour rechercher activement les menaces et les indicateurs d'attaque dans le réseau, permettant ainsi une détection plus rapide.

- Formation et sensibilisation : La formation régulière du personnel sur les nouvelles menaces et les meilleures pratiques de cybersécurité peut améliorer la préparation et la réactivité de l'équipe face aux attaques.

- Déploiement de solutions de sécurité avancées : L'utilisation de technologies de détection avancées, telles que l'IA et le Machine Learning, peut aider à identifier les comportements malveillants et à détecter les menaces plus rapidement.

En combinant ces mesures avec une approche proactive de la cybersécurité, les organisations peuvent mieux se préparer à faire face aux menaces émergentes et améliorer leur capacité à répondre rapidement et efficacement aux incidents de sécurité.

2.2.5 Identification et anticipation des nouvelles menaces

L'identification et l'anticipation des nouvelles menaces sont cruciales pour améliorer les délais de réponse en cas de cyberattaque. Les cybercriminels cherchent constamment à développer de nouvelles techniques et à exploiter des vulnérabilités inconnues pour contourner les mesures de sécurité existantes. Cela rend d'autant plus difficile pour les équipes de cybersécurité de détecter et de répondre rapidement aux attaques.

Comme points importants concernant l'identification et l'anticipation des nouvelles menaces nous pouvons citer :

- Menaces inconnues : Les attaques zero-day, qui exploitent des vulnérabilités non corrigées et inconnues, peuvent être particulièrement dangereuses car elles ne sont pas détectées par les systèmes de sécurité traditionnels. Les équipes de cybersécurité doivent être conscientes de ces menaces potentielles et prendre des mesures pour les anticiper et les contrer.

- Détection comportementale : En utilisant des techniques de détection basées sur le comportement, comme l'analyse des schémas d'activité et l'identification des comportements anormaux, les équipes de sécurité peuvent mieux détecter les attaques qui ne correspondent pas aux signatures connues.

- Surveillance proactive : Les équipes de cybersécurité peuvent adopter une approche proactive de la surveillance du réseau et des systèmes pour détecter les signes précurseurs d'une attaque potentielle.

- Threat Intelligence : L'intégration de Threat Intelligence provenant de sources externes peut aider à identifier les nouvelles menaces et les schémas d'attaque émergents, permettant ainsi aux équipes de mieux se préparer et de répondre rapidement.

- Utilisation du Machine Learning : Le Machine Learning peut être utilisé pour analyser de vastes quantités de données de sécurité, identifier des tendances et des schémas inconnus, et aider à anticiper les attaques futures en se basant sur des comportements d'attaques antérieures.

- Tests de pénétration : Réaliser régulièrement des tests de pénétration et des exercices de simulation d'attaque permet aux équipes de cybersécurité d'évaluer la résilience de leurs systèmes et d'identifier les vulnérabilités potentielles.

La programmation du Machine Learning peut aider à mettre en évidence les points communs entre la nouvelle et les menaces précédemment identifiées lors d'une attaque. L'exécution rapide de cette tâche n'est possible pour aucune équipe. Ce qui met en lumière le côté indispensable des modèles de sécurité adaptative. De ce point de vue, le Machine Learning peut simplifier l'anticipation de nouvelles menaces et réduire le retard grâce à une sensibilisation accrue.

2.2.6 Capacité du personnel

La capacité du personnel fait partie des problèmes courants auxquels font face un grand nombre d'équipes informatiques et de cybersécurité. En fonction des besoins d'une entreprise, le nombre de professionnels qualifiés peut être limité.

Le plus souvent, faire appel à des spécialistes est parfois trop couteux pour les entreprises. Le personnel doit non seulement être rémunéré pour son travail quotidien, mais également pouvoir suivre des formations et obtenir des certifications. Il faut rester au fait des exigences des professionnels de la cybersécurité, notamment concernant l'innovation permanente.

Les outils de sécurité basés sur l'IA peuvent permettre de réduire le personnel de l'équipe tout en augmentant son efficacité.

2.2.7 Adaptabilité

Ce facteur est assez différent des autres déjà évoqués mais est aussi d'une grande importance. L'adaptabilité permet à une équipe ou une organisation de gagner beaucoup de temps. Nous pouvons prendre l'exemple d'une équipe qui doit utiliser une nouvelle solution de sécurité pour l'entreprise. La durée de l'installation à la formation du personnel pour l'utilisation de cet outil de sécurité peut être largement réduite grâce à l'utilisation d'un algorithme de ML qui fait un apprentissage extrêmement rapide avec la puissance des machines de nouvelle génération.

2.3 Utilisation de l'IA dans la cybersécurité

L'intelligence artificielle (IA) dans le domaine de la cybersécurité est en effet considérée comme un super ensemble de disciplines, car elle englobe un large éventail de techniques et d'approches, y compris celles basées sur le Machine Learning et le Deep Learning. Cependant, l'IA ne joue généralement pas un rôle autonome, mais plutôt un rôle complémentaire aux autres méthodes de cybersécurité. Voici quelques raisons pour lesquelles l'IA dans la cybersécurité est considérée comme un complément plutôt qu'une solution autonome :

- Expertise humaine : Malgré les avancées de l'IA, l'expertise humaine reste essentielle en cybersécurité. Les équipes de sécurité doivent toujours jouer un rôle clé dans l'analyse des menaces, la prise de décisions stratégiques et la mise en place de politiques de sécurité.

- Compréhension du contexte : L'IA peut aider à analyser de grandes quantités de données, mais elle peut parfois manquer de compréhension contextuelle. Les humains peuvent apporter une compréhension approfondie du contexte de l'organisation, ce qui est essentiel pour interpréter correctement les résultats de l'IA.

- Détection de menaces inconnues : Bien que l'IA puisse être efficace pour détecter des schémas de comportement malveillant connus, elle peut être moins efficace pour détecter des menaces inconnues et innovantes. C'est là que l'expertise humaine et la pensée créative sont nécessaires.

- Validation des résultats : Les résultats de l'IA doivent souvent être validés et vérifiés par des experts en cybersécurité pour s'assurer de leur pertinence et de leur précision.

- Adaptabilité : Les cybermenaces évoluent rapidement, et il est essentiel d'adapter en permanence les stratégies de cybersécurité pour rester efficace. L'IA peut contribuer à cette adaptabilité, mais elle doit être complétée par des efforts humains pour suivre les évolutions du paysage des menaces.

Le rôle idéal de l'IA en matière de cybersécurité réside dans l'interprétation des schémas établis par les algorithmes de Machine Learning. Naturellement, il n'est pas encore possible pour l'IA d'interpréter les résultats avec les capacités d'un être humain. Les chercheurs s'efforcent de développer ce domaine pour reproduire le fonctionnement de l'intelligence humaine, mais l'IA est un objectif distant qui nécessite que les machines utilisent des concepts abstraits de situations pour les restructurer. En d'autres termes, ce niveau de créativité et de pensée critique n'est pas aussi proche que les rumeurs sur l'IA le laissent entendre.

2.3.1 Comment le Machine Learning est-il utilisé en cybersécurité ?

Les solutions de sécurité basées sur le Machine Learning diffèrent de celles que nous pensons être fournies par l'intelligence artificielle mais elles sont de loin les outils d'IA les plus efficaces que nous avons en matière de cybersécurité. Dans le cadre de cette technologie, le traitement des données est plutôt orienté vers la probabilité qu'un événement se produise ou non.

L'IA a pour objectif d'obtenir des résultats avec succès alors que le but du ML est l'exactitude. Cela signifie qu'il vise à en savoir le plus possible sur un ensemble de données orienté tâche. Il obtient ainsi les performances optimales d'une tâche donnée. Il s'efforce d'atteindre l'unique solution possible à partir des données recueillies, même si elle n'est pas idéale. Avec le ML, il n'existe aucune véritable interprétation des données, ce qui signifie que la responsabilité incombe toujours aux humains.

Le Machine Learning excelle dans les tâches pénibles telles que l'identification et l'adaptation des schémas de données. Ce type de tâche ne sont pas faciles à assurer par l'humain car nécessitent beaucoup de temps, de concentration et d'énergie. Ainsi, même si l'interprétation

de l'analyse des données reste aux mains des humains, le Machine Learning contribue à structurer les données pour obtenir une présentation lisible et prête à être analysée à travers un reporting précis. La cybersécurité basée sur le Machine Learning existe sous plusieurs formes, chacune présentant des avantages spécifiques

2.3.2 Classification des données

La classification des données joue un rôle crucial dans le croisement entre le Machine Learning et la cybersécurité. Elle consiste à attribuer des catégories ou des étiquettes prédéfinies à des points de données en utilisant des règles et des modèles établis. Cette étape est essentielle pour créer des profils d'attaques, de vulnérabilités et d'autres aspects liés à la sécurité proactive. Voici comment la classification des données se situe à l'intersection du Machine Learning et de la cybersécurité :

- Création de profils de menaces : En utilisant des modèles de classification, les données provenant d'attaques passées, de comportements malveillants et d'autres sources de menaces peuvent être étiquetées et regroupées en catégories spécifiques. Ces profils aident les équipes de sécurité à mieux comprendre les méthodes et les schémas utilisés par les attaquants.

- Détection d'anomalies : La classification des données peut être utilisée pour identifier des anomalies dans le comportement normal du réseau ou des utilisateurs. En classant les activités en normales et anormales, il devient possible de détecter rapidement les comportements suspects.

- Création de signatures : Dans le domaine de la cybersécurité, les signatures sont des modèles qui correspondent à des attaques spécifiques. La classification des données permet de créer des signatures en identifiant les caractéristiques distinctives des attaques connues.

- Prévention des intrusions : Les systèmes de prévention des intrusions (IPS) et les systèmes de détection d'intrusion (IDS) utilisent souvent la classification des données pour identifier les comportements ou les trafics malveillants, en comparant les données en temps réel avec des modèles préétablis.

- Apprentissage supervisé : La classification des données est un exemple d'apprentissage supervisé en Machine Learning. Les modèles sont formés sur des données étiquetées, ce qui permet au système d'apprendre à catégoriser correctement de nouvelles données non étiquetées.

2.3.3 Mise en cluster des données

La mise en cluster des données est une technique utile en cybersécurité pour regrouper des données similaires ou inhabituelles dans des ensembles de données appelés "clusters". Cette technique permet d'identifier des motifs, des comportements anormaux et des caractéristiques communes au sein de ces groupes de données. La mise en cluster des données est utilisée dans le contexte de la cybersécurité par :

- Détection d'anomalies : La mise en cluster peut être utilisée pour détecter des anomalies dans le comportement du réseau ou des utilisateurs. Les données qui ne

s'intègrent pas bien dans les clusters normaux peuvent être considérées comme des comportements suspects et potentiellement malveillants.

- Analyse des attaques inconnues : Lorsqu'un système de cybersécurité n'a pas encore été formé pour détecter certaines attaques, la mise en cluster peut être utilisée pour regrouper des schémas de données inconnus ou atypiques. Cela peut aider les analystes de sécurité à comprendre comment une attaque s'est produite et à identifier les éléments exploités ou exposés.

- Profilage des menaces : En examinant les caractéristiques des clusters d'attaques ou de comportements malveillants, les équipes de cybersécurité peuvent créer des profils de menaces plus précis. Ces profils peuvent être utilisés pour détecter de manière proactive des comportements similaires à l'avenir.

- Découverte de nouvelles vulnérabilités : Les clusters peuvent révéler des schémas de comportement inhabituels qui pourraient indiquer de nouvelles vulnérabilités ou des points d'entrée potentiels pour les attaques.

- Apprentissage non supervisé : La mise en cluster est une technique d'apprentissage non supervisé, ce qui signifie qu'elle peut être utilisée pour identifier des structures dans les données sans nécessiter d'étiquettes préalables.

Cependant, il est important de noter que la mise en cluster des données peut également générer des résultats erronés ou non pertinents si elle est mal utilisée. Une compréhension approfondie des données et des techniques de mise en cluster est nécessaire pour interpréter correctement les résultats et éviter les faux positifs ou les faux négatifs.

2.3.4 Actions recommandées

Les actions recommandées sont des conseils et des suggestions basés sur des schémas comportementaux appris par le système et sur les décisions prises dans le passé. Les quelques points clés à retenir concernant les actions recommandées dans ce contexte sont :

- Mesures proactives : Les actions recommandées visent à renforcer les mesures proactives de sécurité en identifiant les comportements potentiellement malveillants et en suggérant des actions pour les contrer avant qu'une menace ne se concrétise.

- Schémas comportementaux : Les actions recommandées reposent sur la détection de schémas et de comportements anormaux ou suspects. L'IA apprend à partir de données historiques pour repérer ces modèles et anticiper les activités malveillantes futures.

- Anciennes décisions : Les actions recommandées tiennent compte des décisions prises dans le passé en réponse à des situations similaires. L'IA peut utiliser ces décisions pour générer des recommandations pertinentes dans des scénarios similaires.

- Structure de conclusion adaptative : L'IA établit des relations logiques en utilisant des points de données préexistants. Cela lui permet de déterminer les actions recommandées en fonction de la situation actuelle, en se basant sur des conclusions adaptatives plutôt que sur des règles strictes et prédéfinies.

- Aucune décision autonome : Il est important de souligner que les actions recommandées ne prennent pas de décisions autonomes. Elles fournissent des suggestions et des conseils aux équipes de sécurité, qui demeurent responsables de l'évaluation finale et de la prise de décision.

- Contribution à la réponse aux menaces : Les actions recommandées contribuent à répondre aux menaces de manière proactive en fournissant aux équipes de sécurité des informations exploitables pour prendre des mesures rapidement et efficacement.

Cependant, ces recommandations doivent toujours être évaluées et validées par des professionnels avant d'être mises en œuvre, car la prise de décision humaine reste essentielle pour garantir des réponses appropriées et adaptées au contexte.

2.3.5 Synthèse des possibilités

Il s'agit d'une étape importante où l'IA analyse les données antérieures, les modèles comportementaux et les nouvelles informations pour identifier de nouvelles opportunités et possibilités. Nous trouvons quelques points clés à retenir concernant la synthèse des possibilités qui sont :

- Nouvelles opportunités : La synthèse des possibilités vise à identifier de nouvelles occasions ou perspectives à partir des données historiques et des nouveaux ensembles de données. Elle vise à tirer parti des leçons apprises pour améliorer la détection des menaces et la gestion des risques.

- Comparaison aux situations antérieures : Contrairement aux recommandations qui fournissent des conseils spécifiques, la synthèse des possibilités cherche à déterminer les chances qu'une situation actuelle corresponde à des situations similaires du passé. Cela peut aider à anticiper les tendances ou les comportements émergents.

- Analyse préventive : La synthèse des possibilités peut être utilisée pour effectuer une analyse préventive des vulnérabilités et des points faibles des systèmes. En identifiant les schémas qui ont conduit à des problèmes de sécurité dans le passé, l'IA peut aider à prendre des mesures pour atténuer les risques futurs.

- Utilisation de données inconnues : La synthèse des possibilités s'appuie sur des ensembles de données antérieurs, mais elle peut également incorporer des données nouvelles et inconnues pour identifier des tendances ou des corrélations non encore découvertes.

- Prise de décision humaine : Tout comme les actions recommandées, la synthèse des possibilités ne prend pas de décisions autonomes. Elle fournit des informations pour aider les professionnels de la cybersécurité à prendre des décisions éclairées et stratégiques.

- Amélioration continue : En analysant en permanence les données antérieures et les nouvelles informations, la synthèse des possibilités contribue à l'amélioration continue des stratégies de cybersécurité et des plans de prévention.

2.3.6 Prévision prédictive

La prévision prédictive est en effet l'une des fonctionnalités les plus visionnaires du Machine Learning (ML) et elle joue un rôle crucial dans divers processus, y compris ceux liés à la cybersécurité. Comme points importants nous pouvons retenir :

- Anticiper les résultats potentiels : La prévision prédictive utilise des modèles et des algorithmes d'apprentissage automatique pour anticiper les résultats futurs en se basant sur l'évaluation des ensembles de données existants. Cela permet d'identifier les tendances, les comportements et les schémas susceptibles de se produire à l'avenir.

- Modèles de menaces : Dans le domaine de la cybersécurité, la prévision prédictive peut être utilisée pour créer des modèles de menaces. Elle permet de prédire les types d'attaques probables et les vulnérabilités potentielles auxquelles une organisation pourrait être confrontée.

- Prévention des fraudes : La prévision prédictive peut aider à détecter les activités frauduleuses en examinant les schémas de transactions et en identifiant les comportements anormaux associés à la fraude.

- Protection contre les violations de données : En analysant les données de sécurité, la prévision prédictive peut aider à identifier les vulnérabilités potentielles dans les systèmes et les réseaux, ce qui permet de prendre des mesures préventives pour éviter les violations de données.

- Terminal prédictif : La prévision prédictive peut être intégrée dans des terminaux prédictifs, où elle peut contribuer à améliorer la prise de décision en fournissant des informations sur les risques futurs.

- Amélioration continue : La prévision prédictive s'appuie sur l'analyse constante des données et l'ajustement des modèles au fil du temps, ce qui permet une amélioration continue des prévisions.

- Base de l'analyse prédictive : La prévision prédictive constitue la base de l'analyse prédictive plus large, qui vise à utiliser des modèles pour anticiper les résultats et les comportements futurs dans divers domaines

2.4 Exemples d'utilisation du Machine Learning dans le domaine de la cybersécurité

Pour mieux comprendre, voici quelques exemples montrant la valeur du Machine Learning dans le domaine de la cybersécurité :

2.4.1 Classification et conformité de la confidentialité des données

La classification des données consiste à attribuer des niveaux de sensibilité ou de confidentialité aux différents types d'informations traitées au sein d'une organisation. Cette classification permet de distinguer les données en fonction de leur importance, de leur niveau de confidentialité et du risque potentiel associé à leur exposition ou à leur compromission. Les niveaux de classification varient d'une organisation à l'autre, mais ils comprennent généralement des catégories telles que "publique", "interne", "confidentielle" et

"très confidentielle". La classification facilite la mise en place de mesures de sécurité appropriées en fonction du niveau de sensibilité de chaque catégorie de données.

La conformité de la confidentialité des données fait référence à la conformité d'une organisation aux réglementations et normes de protection des données en vigueur. Les réglementations telles que le Règlement général sur la protection des données (RGPD) en Europe et la Health Insurance Portability and Accountability Act (HIPAA) aux États-Unis imposent des exigences strictes pour la collecte, le stockage, le traitement et la transmission des données personnelles et sensibles. Les organisations doivent mettre en place des politiques, des procédures et des mesures techniques pour s'assurer que les données sont traitées conformément à ces réglementations et que la confidentialité des données est maintenue. La combinaison de la classification des données et de la conformité de la confidentialité des données permet aux organisations de :

- Identifier et prioriser les données sensibles : En classifiant les données, les organisations peuvent concentrer leurs efforts de sécurité sur les informations les plus importantes et sensibles.

- Mettre en place des contrôles appropriés : En fonction de la classification des données, des contrôles de sécurité adaptés peuvent être mis en place pour protéger les données contre les menaces et les risques potentiels.

- Répondre aux exigences réglementaires : En se conformant aux réglementations de protection des données, les organisations évitent les sanctions et les amendes potentielles tout en renforçant la confiance des clients et des partenaires.

- Gérer les risques : La classification et la conformité aident à identifier et à gérer les risques liés à la confidentialité des données, en prenant des mesures préventives pour éviter les violations de données.

2.4.2 Profils de sécurité des comportements utilisateurs

Ces profils visent à identifier les schémas de comportement typiques des utilisateurs afin de détecter rapidement les activités anormales, potentiellement malveillantes ou non conformes. Parmi les points essentiels nous pouvons citer :

- Création de modèles : Les profils de sécurité des comportements utilisateurs sont créés en collectant et en analysant des données sur les activités et les interactions des utilisateurs. Cela peut inclure des informations telles que les heures de connexion, les types d'actions effectuées, les ressources accédées et les motifs de navigation.

- Schémas de comportement : En examinant les données sur une période de temps, des schémas comportementaux typiques peuvent être identifiés pour chaque utilisateur. Ces schémas peuvent varier en fonction du rôle de l'utilisateur, de ses habitudes et de ses responsabilités.

- Détection des anomalies : Une fois que les schémas comportementaux normaux sont établis, les activités anormales peuvent être détectées en comparant le comportement actuel d'un utilisateur à son modèle habituel. Les déviations significatives peuvent indiquer une utilisation inappropriée, une activité malveillante ou une compromission de compte.

🕮 Adaptabilité : Les profils de sécurité des comportements utilisateurs sont adaptables et doivent être mis à jour en fonction des changements de comportement ou des nouvelles activités des utilisateurs. Cela permet de réduire les faux positifs et d'améliorer la précision de la détection des anomalies.

🕮 Réponse rapide : En détectant rapidement les comportements anormaux, les organisations peuvent réagir plus rapidement aux menaces potentielles et prendre des mesures pour atténuer les risques.

🕮 Intégration avec l'IA : Les techniques d'apprentissage automatique, telles que le Machine Learning, peuvent être utilisées pour améliorer la création de modèles et la détection des anomalies. Les algorithmes peuvent identifier des schémas subtils difficiles à repérer manuellement.

🕮 Respect de la vie privée : La mise en place de profils de sécurité des comportements utilisateurs nécessite une gestion appropriée de la vie privée et des données personnelles. Les informations collectées doivent être protégées et utilisées conformément aux réglementations en vigueur.

2.4.3 Profils de sécurité des performances système

Le concept de profil de diagnostic personnalisé des performances d'un ordinateur est similaire à celui des profils de comportements utilisateurs en cybersécurité. Il s'agit d'établir des modèles de fonctionnement typiques pour les performances de l'ordinateur dans son ensemble, afin de détecter les anomalies potentielles et les signes d'activités malveillantes. Voici quelques points importants à retenir concernant les profils de diagnostic personnalisé des performances :

🕮 Surveiller les performances : En surveillant les performances de l'ordinateur, on peut obtenir des données sur l'utilisation du processeur, de la mémoire, de la bande passante Internet, et d'autres ressources. Les schémas typiques d'utilisation peuvent être établis pour différentes activités.

🕮 Détection des anomalies : Une fois que les schémas de performances typiques sont établis, toute déviation significative peut être considérée comme une anomalie potentielle. Cela pourrait indiquer une activité malveillante ou des problèmes de performance.

🕮 Différenciation des comportements normaux : Il est important de prendre en compte les comportements normaux qui peuvent entraîner des pics d'utilisation de ressources. Par exemple, des téléchargements de fichiers volumineux pourraient être une activité courante pour certains utilisateurs.

🕮 Identification des activités malveillantes : Des indicateurs tels qu'une utilisation excessive de la mémoire ou une utilisation inattendue du processeur pourraient signaler des activités malveillantes, telles que l'exécution de logiciels malveillants en arrière-plan.

🕮 Utilisation de seuils : Pour détecter les anomalies, des seuils peuvent être définis pour différents paramètres de performance. Si ces seuils sont dépassés, une alerte peut être générée pour signaler une activité suspecte.

- Comparaison aux schémas historiques : Les données historiques de performances peuvent être utilisées pour comparer les schémas actuels et antérieurs, ce qui permet de détecter les changements et les variations.

- Amélioration avec l'IA : L'utilisation de l'apprentissage automatique peut permettre d'identifier des anomalies subtiles et complexes dans les performances de l'ordinateur.

2.4.4 Blocage des robots basés sur le comportement

Le blocage des robots basés sur le comportement en cybersécurité fait référence à la mise en place de mesures pour identifier et bloquer les activités des robots ou des bots malveillants en se basant sur leurs schémas de comportement. Les bots sont des programmes informatiques automatisés qui peuvent être utilisés pour diverses activités, allant de l'exploration de sites Web à des activités malveillantes telles que le vol d'informations ou les attaques de déni de service. Voici comment le blocage des robots basés sur le comportement se fait :

- Surveillance du comportement : Les bots malveillants ont souvent des comportements caractéristiques, tels que l'accès fréquent à certaines pages, des taux de clics anormalement élevés, ou des modèles d'interaction prévisibles. Les systèmes de sécurité surveillent ces comportements pour détecter les activités suspectes.

- Analyse des schémas : En collectant des données sur les activités des bots au fil du temps, les systèmes de sécurité peuvent identifier les schémas comportementaux typiques. Cela peut inclure la fréquence des requêtes, les horaires d'activité, les séquences d'actions, etc.

- Détection des anomalies : Les systèmes de sécurité utilisent des algorithmes pour comparer le comportement actuel des bots avec les schémas comportementaux normaux. Si des anomalies significatives sont détectées, cela peut indiquer qu'un bot malveillant est en action.

- Mise en quarantaine ou blocage : Lorsqu'un comportement suspect est identifié, les systèmes de sécurité peuvent prendre des mesures pour mettre en quarantaine ou bloquer les bots malveillants. Cela peut se faire en limitant l'accès aux ressources, en ajoutant des captchas, ou en utilisant des listes noires d'adresses IP connues pour être associées à des activités malveillantes.

- Apprentissage automatique : Les techniques d'apprentissage automatique peuvent être utilisées pour améliorer la détection en identifiant des schémas comportementaux subtils ou en constante évolution.

- Adaptabilité : Les bots malveillants peuvent changer de comportement pour éviter la détection. Les systèmes de sécurité doivent être adaptables et capables de s'ajuster aux nouveaux schémas de comportement[7].

2.5 Conclusion :

[7] https://www.kaspersky.fr/resource-center/definitions/ai-cybersecurity Blog Kaspersky

L'intelligence Artificielle est un domaine qui demeure toujours nouveau car ayant beaucoup de pistes de découverte et la maîtrise de sa capacité n'est pas prêt d'être acquise. L'utiliser pour augmenter la capacité des systèmes à assurer la sécurité peut être d'un grand atout mais présente aussi des risques car l'IA est aussi utilisée par les cybercriminels et est parfois dupée, ce qui devient un point d'attention très important à prendre en compte pour pouvoir assurer la sécurité des systèmes d'information.

CHAPITRE 3 :

Early Warning Detection ET IA

Dans ce chapitre, nous allons élaborer l'utilité de la corrélation entre un système d'alerte précoce et l'intelligence artificielle. Ce travail se fera en premier en exposant différents algorithmes de Machine Learning avec leur utilité et cas d'utilisation, et en second plan nous feront une partie pratique en utilisant un de ces algorithmes pour la détection d'une menace.

3.1 Définition

Un système d'alerte précoce pour la cybersécurité utilisant un algorithme de Machine Learning est un système automatisé conçu pour détecter et signaler rapidement les menaces et les anomalies potentielles dans les réseaux informatiques et les systèmes d'information. C'est un ensemble de processus, de technologies et d'algorithmes de Machine Learning utilisés pour surveiller en temps réel le trafic réseau, le comportement des utilisateurs, les journaux d'événements et d'autres données pertinentes liées à la sécurité informatique. L'objectif principal est d'identifier les menaces potentielles le plus tôt possible, avant qu'elles ne causent des dommages ou des perturbations majeures.

3.2 Algorithme de Machine Learning Utilisé par les EWS

Plusieurs algorithmes de ML sont utilisés par les EWS pour faciliter les processus d'analyse, de détection des menaces, de corrélation des évènements et de prise de décision lors d'une cyberattaque. Parmi ces algorithmes nous pouvons citer :

- K-Nearest Neighbors (KNN) : K-Nearest Neighbors (KNN) peut être utilisé pour la détection de menaces en cybersécurité, mais son application dépend du contexte spécifique et de la manière dont il est configuré. Voici comment KNN peut être adapté à la détection de menaces en cybersécurité :

🕮 Collecte de données : Tout d'abord, vous avez besoin d'un ensemble de données d'entraînement qui comprend des exemples de trafic réseau ou de comportement système, étiquetés comme "normal" ou "anormal" en fonction de leur conformité aux modèles de comportement attendus.

🕮 Choix du nombre de voisins (K) : Vous devez décider d'un nombre K approprié, c'est-à-dire combien de voisins les plus proches vous voulez examiner lors de la détection d'une menace. Cette valeur peut varier en fonction de la complexité de votre système et de vos besoins de sécurité.

🕮 Mesure de la distance : Pour détecter des menaces, vous devez définir une mesure de distance appropriée pour évaluer la similarité entre les comportements. Cela peut être basé sur des caractéristiques telles que les protocoles réseau utilisés, les ports, les fréquences, etc.

🕮 Modélisation : En utilisant votre ensemble de données d'entraînement, KNN peut être utilisé pour créer un modèle de comportement "normal" en classifiant les données comme "normal" ou "anormal" en fonction des K voisins les plus proches.

🕮 Détection d'anomalies : Lorsque vous recevez de nouvelles données en temps réel, vous pouvez utiliser votre modèle KNN pour classer le comportement comme "normal" ou

"anormal". Si un comportement est considéré comme anormal (c'est-à-dire qu'il est éloigné des comportements "normaux" du modèle), il peut être identifié comme une menace potentielle.

- Réglage de seuil : Vous devrez définir un seuil pour décider quand un comportement est considéré comme "anormal". Cela peut être basé sur des mesures telles que la distance moyenne des K voisins les plus proches.

- Mise à jour du modèle : Les modèles KNN peuvent nécessiter une mise à jour régulière pour s'adapter aux évolutions du trafic réseau ou du comportement système. Les menaces évoluent, et votre modèle doit suivre le rythme.

KNN peut être efficace pour la détection d'anomalies simples dans des environnements de cybersécurité, mais il peut avoir des limitations en termes de performances et de capacité à gérer des données en temps réel à grande échelle. Dans des environnements de sécurité plus complexes, d'autres techniques d'apprentissage automatique et d'analyse de données, telles que les réseaux de neurones, les méthodes basées sur les arbres de décision ou les approches basées sur les signatures, peuvent être nécessaires pour une détection plus précise et rapide des menaces.

- Random Forest :

Random Forest est un algorithme d'ensemble puissant en apprentissage automatique qui est largement utilisé dans divers domaines, y compris la cybersécurité, pour la détection d'intrusions et la classification des événements de sécurité. Ainsi fonctionne le Random Forest :

- Construction des arbres de décision : Le Random Forest commence par créer un ensemble d'arbres de décision. Chaque arbre est construit à partir d'un échantillon aléatoire de l'ensemble de données d'entraînement, ce qui contribue à l'indépendance entre les arbres.

- Bagging (Bootstrap Aggregating) : L'algorithme utilise une technique appelée "bagging" pour échantillonner aléatoirement les données avec remplacement, ce qui signifie qu'une même observation peut être incluse dans plusieurs arbres. Cela permet d'introduire de la variabilité dans les arbres et contribue à réduire la variance.

- Division des données : Chaque arbre de décision est construit en divisant les données en fonction de caractéristiques spécifiques (par exemple, les attributs du trafic réseau en cybersécurité) pour séparer les classes (par exemple, les événements normaux et les événements suspects).

- Prédiction : Une fois que les arbres sont construits, le Random Forest effectue des prédictions en combinant les résultats de chaque arbre. Pour la classification, il peut s'agir d'un vote majoritaire, où la classe la plus fréquemment prédite par les arbres est attribuée à l'événement.

Avantages en cybersécurité :

- Précision élevée : Le Random Forest est connu pour sa précision en classification, ce qui en fait un choix solide pour la détection d'intrusions. Il peut distinguer efficacement entre les événements normaux et les événements malveillants.

- Gestion des données déséquilibrées : En cybersécurité, il est courant d'avoir un déséquilibre entre les classes (il y a généralement plus d'événements normaux que d'événements malveillants). Le Random Forest peut traiter ce déséquilibre de manière efficace.

- Capacité à gérer de nombreuses fonctionnalités : En cybersécurité, il peut y avoir de nombreuses caractéristiques à prendre en compte, telles que les adresses IP, les ports, les protocoles, etc. Le Random Forest peut gérer de manière efficace des ensembles de données riches en fonctionnalités.

- Robustesse aux données bruitées : Il est robuste aux données bruitées et aux valeurs aberrantes, ce qui est important en cybersécurité, car les données peuvent être incomplètes ou corrompues.

En résumé, le Random Forest est un algorithme très précieux en cybersécurité grâce à sa capacité à fournir des prédictions de haute qualité, à gérer les données déséquilibrées et à gérer de nombreuses fonctionnalités. Il est largement utilisé pour la détection d'intrusions et la classification des événements de sécurité.

- Naive Bayes :

L'algorithme Naive Bayes est effectivement un outil populaire en apprentissage automatique pour la classification, y compris la classification des événements de sécurité en cybersécurité.

Fonctionnement du Naive Bayes :

- Théorème de Bayes : Le Naive Bayes est basé sur le théorème de Bayes, qui traite des probabilités conditionnelles. Il permet de calculer la probabilité d'une classe (par exemple, "événement malveillant" ou "événement normal") étant donné certaines caractéristiques ou attributs (par exemple, les caractéristiques d'un événement de sécurité).

- Assomption naïve : L'assomption "naïve" du Naive Bayes réside dans le fait qu'il suppose que toutes les caractéristiques sont indépendantes les unes des autres, ce qui simplifie les calculs. En réalité, cette supposition peut ne pas être tout à fait vraie, mais elle permet des calculs rapides et efficaces.

- Calcul des probabilités : Le Naive Bayes calcule la probabilité a posteriori de chaque classe étant donné les valeurs observées des caractéristiques. Il utilise les probabilités conditionnelles et les probabilités a priori pour effectuer ces calculs.

- Classification : Une fois que les probabilités de chaque classe sont calculées, le Naive Bayes attribue l'observation à la classe avec la probabilité la plus élevée. Par exemple, si la probabilité d'un événement étant malveillant est plus élevée que la probabilité d'être normal, il sera classifié comme malveillant.

Avantages en cybersécurité :

- Efficace pour des problèmes de classification à grande échelle : Le Naive Bayes est rapide et efficace, ce qui en fait un choix approprié pour la classification d'événements de sécurité à grande échelle, où de nombreuses observations doivent être traitées rapidement.

- Bonne performance avec des ensembles de données de taille modeste : Il fonctionne bien même avec des ensembles de données relativement petits, ce qui peut être le cas en cybersécurité lorsque les données d'incident spécifiques sont limitées.

- Interprétabilité : Les résultats de l'algorithme sont généralement faciles à interpréter, ce qui est important en cybersécurité pour comprendre pourquoi une décision de classification a été prise.

Il est important de noter que l'algorithme de Naive Bayes repose sur l'assomption naïve d'indépendance des caractéristiques, ce qui peut ne pas être approprié pour tous les types de données en cybersécurité. Dans certains cas, d'autres techniques d'apprentissage automatique, telles que les réseaux de neurones ou les arbres de décision, peuvent être préférées pour une classification plus précise mais l'algorithme reste un outil utile dans l'arsenal de la cybersécurité pour la classification des événements de sécurité.

- Support Vector Machines (SVM) :

Les Support Vector Machines (SVM) sont un autre algorithme d'apprentissage automatique couramment utilisé en cybersécurité, notamment pour la classification binaire, telle que la séparation des événements de sécurité malveillants des événements normaux.

Fonctionnement des Support Vector Machines (SVM) :

- Hyperplan de séparation : Les SVM visent à trouver un hyperplan de séparation optimal dans un espace de caractéristiques. Cet hyperplan est choisi de manière à maximiser la marge entre les deux classes (par exemple, événements malveillants et événements normaux).

- Transformation de l'espace : Si les données ne sont pas linéairement séparables (c'est-à-dire qu'un simple hyperplan ne peut pas séparer les classes), les SVM peuvent utiliser une technique de transformation de l'espace pour projeter les données dans un espace de dimension supérieure où elles sont séparables.

- Vecteurs de support : Les vecteurs de support sont les échantillons d'entraînement qui se trouvent le plus près de l'hyperplan de séparation. Ils jouent un rôle crucial dans le fonctionnement des SVM.

- Fonction de marge : Les SVM utilisent une fonction de marge pour mesurer la distance entre les vecteurs de support et l'hyperplan de séparation. L'objectif est de maximiser cette marge tout en minimisant l'erreur de classification.

- Classification : Une fois que l'hyperplan optimal est trouvé, les SVM peuvent être utilisés pour classer de nouvelles observations en fonction de leur position par rapport à cet hyperplan. Par exemple, si une nouvelle observation se trouve du côté positif de l'hyperplan, elle sera classée comme malveillante, sinon elle sera classée comme normale.

Avantages en cybersécurité :

- Efficacité dans les espaces de grande dimension : Les SVM fonctionnent bien dans des espaces de caractéristiques de grande dimension, ce qui est courant en cybersécurité où de nombreuses caractéristiques (par exemple, les attributs du trafic réseau) doivent être prises en compte.

- Capacité à traiter des données non linéaires : Grâce à la transformation de l'espace, les SVM peuvent traiter des données qui ne sont pas linéairement séparables, ce qui est utile en cybersécurité lorsque les relations entre les caractéristiques ne sont pas simples.

- Bonne généralisation : Les SVM ont tendance à bien généraliser à partir des données d'entraînement, ce qui signifie qu'ils peuvent fonctionner efficacement avec de nouvelles données.

Néanmoins les SVM peuvent devenir complexes à ajuster et à configurer, et leur performance dépend souvent du choix de la fonction de noyau (kernel function) appropriée. En cybersécurité, où la précision et la vitesse de détection sont cruciales, les SVM peuvent être des outils puissants pour la classification des événements de sécurité.

- Réseaux de neurones (Neural Networks) :

Les réseaux de neurones profonds, également connus sous le nom de réseaux de neurones profonds (ou Deep Learning), sont en effet des outils puissants en cybersécurité et dans de nombreux autres domaines.

Fonctionnement des réseaux de neurones profonds :

- Architecture en couches : Les réseaux de neurones profonds sont composés de multiples couches de neurones interconnectées. Chaque couche reçoit des données en entrée, effectue des calculs et transmet les résultats à la couche suivante. Ces couches sont souvent divisées en trois types : la couche d'entrée, les couches cachées et la couche de sortie.

- Apprentissage profond : Ce qui distingue les réseaux de neurones profonds, c'est leur capacité à apprendre des représentations hiérarchiques complexes à partir des données. Chaque couche du réseau extrait des caractéristiques de plus en plus abstraites à mesure que l'on progresse vers les couches supérieures.

- Réseaux de neurones convolutionnels (CNN) : En cybersécurité, les réseaux de neurones convolutionnels sont couramment utilisés pour la classification d'images, la détection d'intrusions basée sur l'analyse de paquets réseau et la détection d'attaques basées sur des données de capteurs.

- Réseaux de neurones récurrents (RNN) : Les RNN sont utilisés pour traiter des données séquentielles, telles que les journaux d'événements. Ils sont capables de modéliser des dépendances temporelles dans les données, ce qui est crucial pour la détection d'intrusions en temps réel.

Avantages en cybersécurité :

- Capacité de détection avancée : Les réseaux de neurones profonds peuvent détecter des modèles complexes dans les données, ce qui est essentiel en cybersécurité pour la détection d'intrusions et la reconnaissance de comportements malveillants.

- Adaptabilité aux données non structurées : En cybersécurité, il est courant de travailler avec des données non structurées, telles que des journaux d'événements. Les réseaux de neurones profonds peuvent apprendre à partir de ces données sans nécessiter une ingénierie des caractéristiques manuelle.

- Amélioration continue : Les réseaux de neurones profonds peuvent s'adapter et s'améliorer avec plus de données d'entraînement, ce qui est essentiel compte tenu de la croissance constante des menaces en cybersécurité.

Cependant, les réseaux de neurones profonds peuvent nécessiter des ensembles de données volumineux et peuvent être exigeants en termes de puissance de calcul. Leur complexité peut également rendre leur interprétation difficile, ce qui est un défi en matière de cybersécurité où la compréhension des décisions prises par le modèle est importante. Malgré cela, les réseaux de neurones profonds sont des outils puissants pour la cybersécurité moderne, en particulier pour des tâches telles que la détection d'intrusions et l'analyse de logs.

- Clustering :

Les algorithmes de clustering, y compris K-Means, sont très importants en cybersécurité pour regrouper des événements de sécurité similaires en fonction de leurs caractéristiques.

Fonctionnement des algorithmes de clustering :

- Groupement par similarité : Les algorithmes de clustering regroupent les données en groupes, appelés clusters, en fonction de la similarité des caractéristiques entre les données. Les données qui sont similaires les unes aux autres sont placées dans le même cluster.

- K-Means : K-Means est l'un des algorithmes de clustering les plus couramment utilisés. Il commence par sélectionner un nombre de clusters (k) et attribue les données aux clusters de manière à minimiser la distance moyenne entre les données et le centre (centroïde) de leur cluster.

- Interprétation des clusters : Une fois que les données sont regroupées en clusters, il est possible d'examiner les caractéristiques des clusters pour identifier des tendances ou des comportements spécifiques.

Avantages en cybersécurité :

- Détection d'anomalies : Les algorithmes de clustering peuvent aider à identifier des comportements anormaux en regroupant les données normales dans un cluster distinct et en identifiant les données qui ne correspondent à aucun cluster.

- Détection de menaces : Ils peuvent être utilisés pour détecter des groupes d'activités malveillantes en regroupant des événements de sécurité liés à une attaque.

- Segmentation des données : Le clustering peut aider à segmenter les données de sécurité en fonction de différents critères, ce qui facilite leur analyse et leur gestion.

- Réduction du bruit : En identifiant les clusters de données normales, les données inhabituelles ou bruyantes peuvent être plus facilement repérées.

Cependant, le choix du nombre de clusters (k) dans l'algorithme K-Means peut être délicat, et il peut être nécessaire d'expérimenter plusieurs valeurs pour trouver la meilleure configuration. De plus, l'interprétation des clusters peut parfois être complexe en fonction de la nature des données en cybersécurité. Néanmoins, le clustering reste un outil important pour la détection d'anomalies et la segmentation des données en cybersécurité.

3.3 Cas d'utilisation

Pour ce cas d'utilisation, nous allons utiliser un algorithme de ML pour la détection de spam. Nous utiliserons le langage Python 3 qui est le langage le plus utilisé en programmation ML grâce à sa maniabilité et sa capacité de chargement et de traitement de données.

Ainsi nous allons utiliser un algorithme tel que le classificateur Naive Bayes, pour détecter les courriers électroniques spam. Voici un exemple simple de programme Python utilisant « scikit-learn », une bibliothèque populaire d'apprentissage automatique, pour créer un modèle de détection de spam.

Nous avons écrit un simple programme adaptable à différents besoins et sources de données.

Nous utilisons quelques exemples de phrases faisant office de petite base de données qui pourrait être remplacée par une grande source de données.

Dans les captures qui suivent nous pourrons voir la déclaration des données et l'utilité de chaque commande.

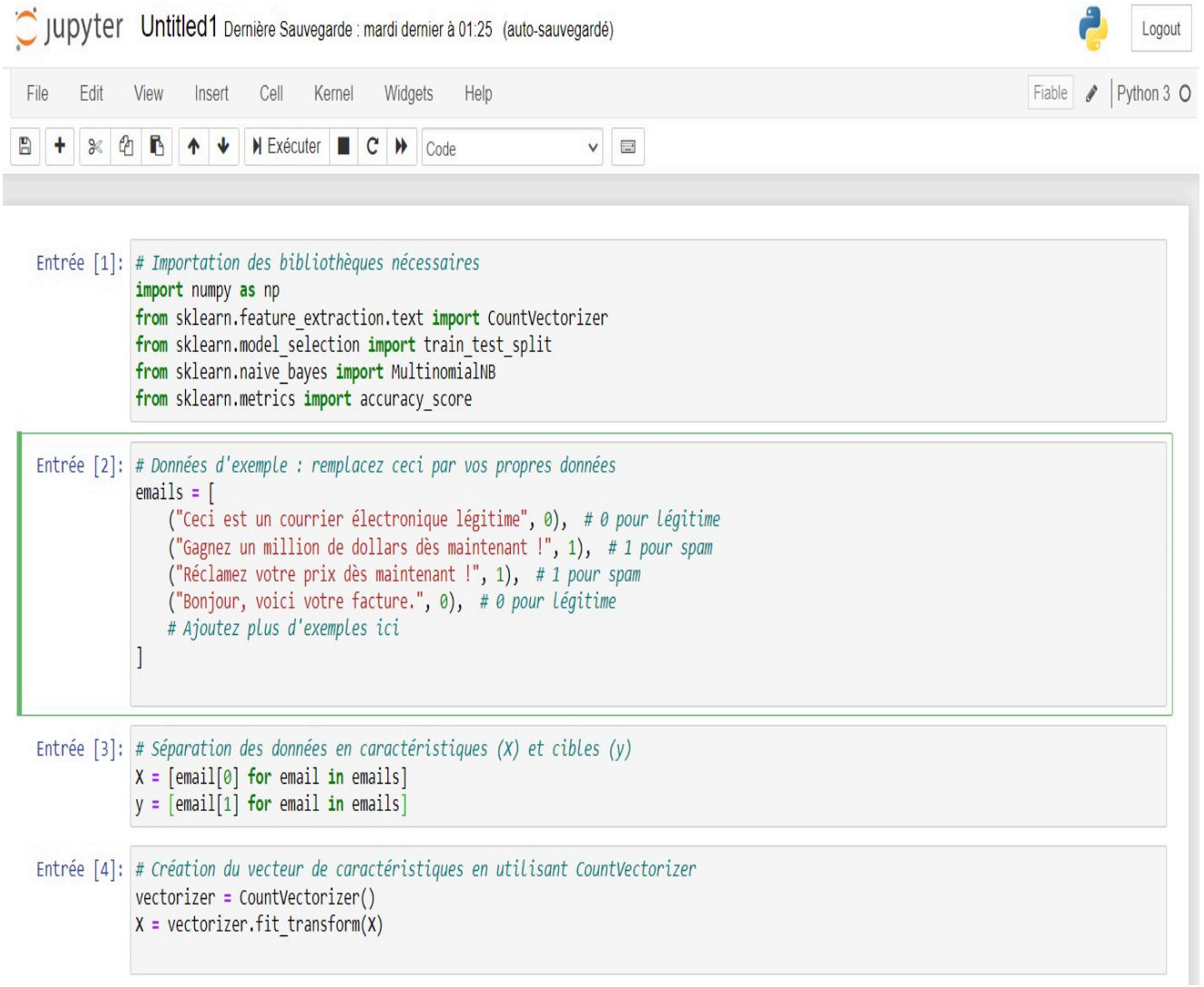


Figure 6: Déclaration des variables de l'algorithme

Dans la seconde capture nous verrons l'entrainement de l'algorithme avec le résultat du test montrant si l'e-mail reçu est un spam ou non.

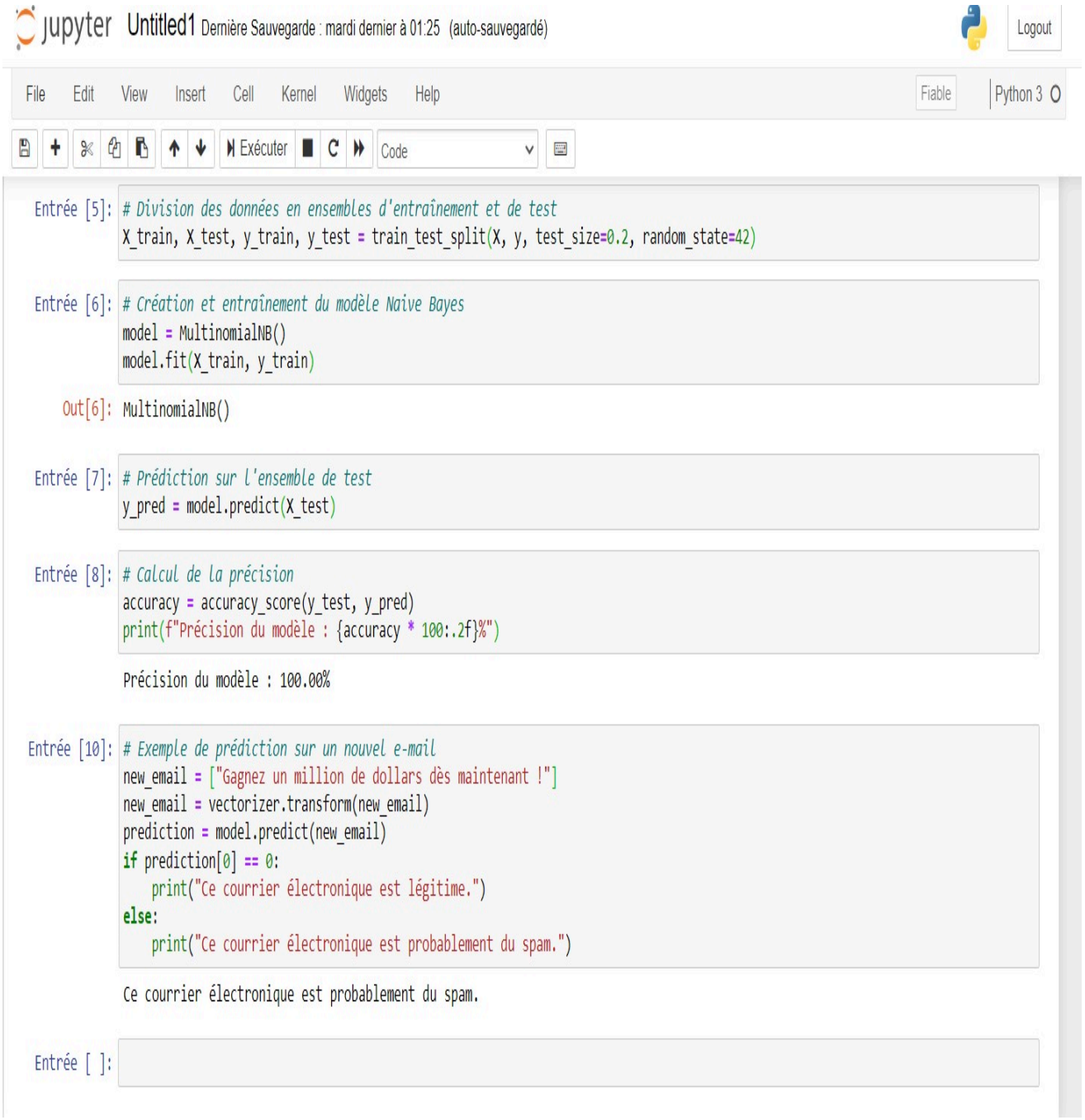


Figure 7: Résultat du test de l'algorithme

Là nous voyons que le courrier électronique reçu et considéré comme un spam.

C'est ce genre d'algorithme qui est utilisé au niveau des SAP pour la détection rapide de spam avec une grande base de données et une configuration optimale des règles car un courrier électronique peut être considéré comme spam par erreur.

3.4 Conclusion

L'utilisation des systèmes d'alerte précoce avec l'intelligence artificielle rend beaucoup plus facile la détection des menaces et la lutte contre les cyberattaques. Bien que cet automatisme ait beaucoup d'avantages, il est aussi plus que nécessaire d'avoir une surveillance humaine pour la prise de décision dans certains cas, surtout ceux avec un faux positif.

Conclusion

Au terme de cette exploration approfondie des Systèmes d'Alerte Précoce (SAP) et de leur convergence avec l'Intelligence Artificielle (IA), il est assez clair que l'union de ces deux domaines offre une réponse puissante et adaptable aux défis complexes de la cybersécurité. Cette conclusion englobe les découvertes et les réflexions acquises tout au long de ce mémoire, mettant en lumière les contributions significatives à notre compréhension de cette dynamique cruciale.

Le premier chapitre a jeté les fondations en définissant les SAP comme des sentinelles numériques, capables de déceler les menaces naissantes avant même qu'elles ne se concrétisent. Nous avons exploré les composants essentiels de ces systèmes, soulignant leur rôle préventif et leur importance croissante dans un paysage numérique en constante évolution.

Le deuxième chapitre a ouvert la boîte noire de l'IA, démontrant comment cette technologie transforme la détection des menaces. Cette convergence des SAP et de l'IA offre plusieurs avantages notables. Elle permet une détection plus rapide et plus précise des menaces, réduisant ainsi le temps de réaction face aux attaques. L'analyse comportementale fine fournie par l'IA offre une capacité d'adaptation aux tactiques évolutives des attaquants, renforçant ainsi la résilience des systèmes de sécurité.

De plus, les algorithmes sophistiqués présentées dans le troisième chapitre, tels que le Random Forest, le Naive Bayes, les Machines à Support Vectoriel, et les Réseaux de Neurones, ont été mis en lumière pour leur capacité à améliorer la précision, à anticiper les anomalies, à évoluer avec les attaques et ont démontré comment cette synergie est déjà mise en œuvre avec succès dans divers contextes, de la détection d'intrusions au filtrage des courriels malveillants, ouvrant ainsi la voie à une sécurité plus robuste.

Cependant, malgré ces succès, des défis subsistent. Les questions éthiques, la confidentialité des données, et la nécessité d'une compréhension approfondie des faux positifs demeurent des préoccupations cruciales. Les futurs travaux devraient être consacrés à la résolution de ces défis, tout en explorant des axes d'amélioration de la collaboration homme-machine dans la gestion des cybermenaces.

En conclusion, ce mémoire a cherché à cartographier une ère nouvelle dans la cybersécurité. Une ère où les SAP alimentés par l'IA ne sont pas simplement des outils, mais des partenaires intelligents dans la défense du cyberespace. Dans un monde où la vitesse des attaques ne cesse d'augmenter, cette alliance offre une défense robuste pour la sécurisation continue de nos systèmes informatiques. Puisse cet étude stimuler de futures avancées, ancrées dans une compréhension approfondie des enjeux et portées par la quête d'une cybersécurité optimale.

Webographie

Steve Morgan, rédacteur en chef Northport, NY – 10 décembre 2022,
https://cybersecurityventures.com/stats/

https://www.websiterating.com/fr/research/cybersecurity-statistics-facts/

Cyber Edu, Forcepoint
https://www.forcepoint.com/fr/cyber-edu/intrusion-prevention-system-ips

https://fr.wikipedia.org/wiki/Syst%C3%A8me_de_d%C3%A9tection_d%27intrusion

Orange Cyberdefence
https://www.orangecyberdefense.com/fr/solutions/protection-des-mobiles-et-des-endpoints/endpoint-detection-and-response-pourquoi-ledr

Blog Microsoft
https://www.microsoft.com/fr/security/business/security-101/what-is-siem#:~:text=Un%20syst%C3%A8me%20de%20gestion%20des,elles%20ne%20perturbent%20leurs%20activit%C3%A9s

https://www.kaspersky.fr/resource-center/definitions/ai-cybersecurity Kaspersky Blog