Ugo BOZION
Thomas GRZESINSKI
04/11/2025

Les types d'IA et leur application à la sécurité (cyber)

Table des matières

La notion d'intelligence artificielle (IA) existe depuis le XX^{e} siècle. En effet, dès les années 1940, les scientifiques américains Warren McCulloch et Walter Pitts ont proposé de reproduire, à l'aide d'une machine, le fonctionnement interne du cerveau humain. Leurs travaux ont conduit à la création du neurone formel, premier modèle mathématique du neurone.

Au fil du temps, l'IA a énormément évolué. Si elle a connu une période de désintérêt dans les années 1970, elle connaît aujourd'hui un essor spectaculaire grâce au deep learning, une technologie qui permet aux machines d'apprendre à partir de grandes quantités de données. Les intelligences artificielles font désormais partie de notre quotidien, même si elles ne sont pas encore capables d'analyser une situation complexe ou de raisonner de manière parfaitement autonome.

De plus, de nombreux modèles d'IA ont vu le jour. Parmi les plus connus, on retrouve ChatGPT, développé par OpenAI, ou encore Apple Intelligence, récemment intégrée aux appareils de la marque à la pomme. D'autres systèmes, comme DeepSeek, ont également suscité un grand intérêt ces dernières années. Ces IA font désormais partie

intégrante de notre vie : par exemple, plus de 2,5 milliards de messages ou requêtes sont envoyés chaque jour à ChatGPT.

Aujourd'hui, nous allons voir quels sont les différents types d'intelligence artificielle existants, ainsi que leurs principales applications dans le domaine de la sécurité.

1 . Les types d'IA

Aujourd'hui, nous interagissons principalement avec les premières itérations des applications d'intelligence artificielle, qui reposent sur des modèles de machine learning.

Le machine learning, ou apprentissage automatique, est un domaine de recherche de l'intelligence artificielle qui vise à donner aux machines la capacité d'apprendre à partir de données sans être explicitement programmées.

Plus précisément, il s'agit d'un procédé par lequel des informations pertinentes sont extraites d'un ensemble de données d'entraînement. L'objectif est de déterminer les paramètres d'un modèle afin qu'il atteigne les meilleures performances possibles lors de la réalisation de la tâche pour laquelle il a été conçu.

Une fois la phase d'apprentissage terminée, le modèle peut être déployé en production, c'est-à-dire utilisé dans des applications réelles.

Un exemple connu d'intelligence artificielle appliquée au grand public est Siri, l'assistant vocal introduit par Apple en 2011 dans son système d'exploitation iOS. À ses débuts, Siri nécessitait encore une forte intervention humaine pour enrichir sa base de connaissances et améliorer ses fonctionnalités.

L'intelligence artificielle reste un domaine en constante évolution. Il est possible de mieux comprendre ses différentes formes en distinguant deux grandes catégories :

- les capacités de l'IA,

- les fonctionnalités de l'IA.

1.1 Trois types d'IA basés sur les capacités

Il existe donc trois types d'IA selon leurs capacités :

- L'intelligence artificielle étroite (ANI, Artificial Narrow Intelligence) :

L'ANI est classée dans la catégorie des IA faibles, car elle est spécialisée dans une gamme étroite de tâches ou de situations, comme la reconnaissance vocale, la traduction automatique ou les voitures autonomes. Par exemple Siri, Alexa ou Google kAssistant sont des intelligences étroites

- L'intelligence artificielle générale (AGI, Artificial General Intelligence) :

L'AGI est considérée comme une IA forte, car elle peut travailler à un niveau comparable à l'intelligence humaine. Elle serait capable de comprendre, d'apprendre et de résoudre des problèmes dans différents domaines, même ceux pour lesquels elle n'a pas été spécifiquement programmée.

- La super-intelligence artificielle (ASI, Artificial Superintelligence) :

Bien que ce type d'IA n'existe pas encore, l'ASI désigne une machine possédant une intelligence supérieure à celle de l'être humain dans tous les domaines, y compris la créativité, le raisonnement et la résolution de problèmes complexes.

1.2 Les types d'IA basés sur la fonctionnalité

L'intelligence artificielle peut également être classée selon sa fonctionnalité, c'est-à-dire ce qu'elle est capable de faire dans la pratique. On distingue quatre types principaux :

- L'intelligence artificielle réactive :

L'IA réactive est composée de systèmes sans mémoire, conçus pour accomplir une tâche très spécifique. Ces machines ne se basent que sur les informations disponibles à un instant donné et ne conservent aucune trace de leurs décisions passées. Elles utilisent des modèles mathématiques et statistiques pour analyser rapidement de grandes quantités de données et produire un résultat pertinent. Parmi les exemples connus, on peut citer IBM Deep Blue, le supercalculateur spécialisé dans le jeu

d'échecs qui a battu Garry Kasparov dans les années 1990, ainsi que les moteurs de recommandation de Netflix, qui analysent l'historique de visionnage pour proposer des contenus adaptés aux utilisateurs.

- L'intelligence artificielle à mémoire limitée :

Contrairement à l'IA réactive, l'IA à mémoire limitée peut se souvenir temporairement d'événements passés et utiliser ces informations pour guider ses décisions futures. Elle observe des objets ou des situations sur une période donnée et améliore ses performances au fil du temps, sans toutefois stocker ces données sur le long terme. Des exemples courants incluent les outils d'IA générative tels que ChatGPT, Bard ou DeepAI, qui prédisent le mot ou l'élément suivant dans un texte ou une image. Les assistants virtuels comme Siri, Alexa ou IBM Watson Assistant utilisent également ce type d'IA pour comprendre les demandes des utilisateurs et fournir des réponses appropriées. Enfin, les voitures autonomes s'appuient sur l'IA à mémoire limitée pour analyser leur environnement en temps réel et prendre des décisions sécurisées de conduite.

- L'intelligence artificielle à théorie de l'esprit :

Ce type d'IA relève de l'intelligence artificielle générale et n'existe pas encore concrètement. L'IA à théorie de l'esprit serait capable de comprendre les émotions, les intentions et les pensées d'autrui, et d'adapter ses interactions en conséquence. Elle pourrait ainsi simuler des relations humaines complexes et contextualiser des informations telles que des œuvres d'art ou des textes littéraires. Un exemple en cours de développement est l'IA émotionnelle, qui analyse des voix, images et autres données pour reconnaître et simuler les émotions humaines. Actuellement, elle n'est pas encore capable de comprendre ou de réagir réellement aux sentiments.

- L'intelligence artificielle consciente :

Enfin, l'IA consciente est strictement théorique. Si elle venait à être réalisée, elle posséderait une conscience de soi, serait capable de comprendre ses propres états internes ainsi que ceux des humains, et pourrait développer ses propres émotions, besoins et croyances. Ce type d'IA correspondrait à la catégorie de la super-intelligence artificielle et demeure pour l'instant un concept futuriste.

2. L'intelligence artificielle et son application dans la cybersécurité

L'intelligence artificielle occupe aujourd'hui une place centrale dans la cybersécurité. Grâce à l'apprentissage automatique et à l'analyse intelligente des données, elle permet d'automatiser la détection des menaces, d'améliorer la réponse aux attaques et de renforcer la protection des systèmes informatiques face à des risques en constante évolution.

L'IA appliquée à la cybersécurité se distingue par sa capacité à analyser d'immenses volumes de données en un temps record, à identifier des schémas inhabituels et à prendre des décisions éclairées sans intervention humaine. Cette rapidité et cette précision dépassent largement les capacités humaines, ce qui en fait un atout essentiel pour les entreprises confrontées à des menaces toujours plus sophistiquées.

2.1 Le rôle de l'IA dans la cybersécurité

Le rôle de l'intelligence artificielle dans la cybersécurité est multiple. Elle automatise des tâches répétitives comme l'analyse des journaux d'activité, la recherche de vulnérabilités ou la surveillance des réseaux. Ainsi, les analystes humains peuvent se concentrer sur des missions plus complexes et stratégiques.

L'IA permet également de détecter les menaces en temps réel, d'identifier des attaques avant même qu'elles ne se produisent, et de réagir instantanément pour les neutraliser. En apprenant continuellement à partir de nouvelles données, elle améliore sans cesse sa capacité à reconnaître et à contrer des attaques émergentes, telles que les ransomwares, le phishing ou les attaques par déni de service (DoS).

2.2 Les avantages de l'IA en cybersécurité

L'utilisation de l'intelligence artificielle dans la cybersécurité présente de nombreux avantages :

- L'automatisation des processus : elle réduit la charge de travail des équipes informatiques en effectuant automatiquement la surveillance, la détection et parfois la réponse aux menaces.

- La réduction des erreurs humaines : contrairement aux systèmes manuels, l'IA limite les risques d'erreur, souvent à l'origine de failles de sécurité.
- Une meilleure prise de décision : les systèmes basés sur l'IA analysent les vulnérabilités, évaluent les priorités et proposent des solutions adaptées pour renforcer la stratégie de sécurité globale.
- La gestion des volumes massifs de données : l'IA peut traiter d'immenses quantités de données issues de multiples sources, permettant d'identifier plus rapidement les anomalies et les signaux faibles souvent invisibles pour l'œil humain.
- L'apprentissage continu : grâce au machine learning et au deep learning, les systèmes d'IA s'améliorent en permanence, rendant leurs défenses de plus en plus robustes face aux menaces nouvelles et imprévisibles.

2.3 Les principales applications de l'IA dans la cybersécurité

L'intelligence artificielle intervient aujourd'hui dans de nombreux domaines de la cybersécurité :

- Protection et authentification : les systèmes d'IA renforcent la sécurité des connexions grâce à des technologies telles que la reconnaissance faciale, les empreintes digitales ou les CAPTCHA intelligents. Ces outils permettent d'empêcher le piratage des mots de passe et les attaques par force brute.

- Détection du phishing : l'IA analyse le contenu et le contexte des e-mails pour identifier les tentatives de phishing, les adresses usurpées ou les noms de domaine falsifiés. Elle apprend également à reconnaître le style et le comportement des utilisateurs afin de détecter les attaques de spear-phishing plus ciblées.
- Gestion des vulnérabilités : grâce à l'analyse comportementale, l'IA identifie les anomalies dans l'activité des serveurs ou des utilisateurs et détecte les menaces de type « zero-day » avant qu'elles ne soient officiellement signalées.

- Sécurité réseau : l'IA contribue à la création de politiques de sécurité intelligentes, à la détection des connexions suspectes et à la mise en place de stratégies de « confiance zéro » (Zero Trust).

- Analyse comportementale : en étudiant les habitudes des utilisateurs et des appareils, l'IA est capable de repérer les comportements inhabituels, signalant ainsi une possible intrusion ou un logiciel malveillant.

2.4 L'IA générative et la cybersécurité

L'IA générative, connue pour sa capacité à produire de nouvelles données à partir de données existantes, représente une avancée majeure dans la cybersécurité moderne. Elle permet de créer des simulations réalistes d'attaques pour tester la résistance des systèmes, de prédire les scénarios futurs en se basant sur les modèles d'attaques passées et d'améliorer la détection grâce à la génération de données synthétiques. Ces applications rendent la défense proactive et permettent aux organisations d'anticiper les menaces avant qu'elles ne surviennent.

2.5 Les limites et les risques de l'IA en cybersécurité

Si l'intelligence artificielle offre des perspectives prometteuses, elle présente également certains risques. Les cybercriminels utilisent eux aussi l'IA pour concevoir des attaques plus complexes, créer des contenus falsifiés (deepfakes) ou contourner les systèmes de défense. De plus, l'automatisation totale comporte un danger : une mauvaise configuration ou un apprentissage biaisé pourrait entraîner une détection erronée et une mauvaise réponse à une menace réelle.

Ainsi, l'intervention humaine demeure essentielle pour superviser, corriger et améliorer continuellement les systèmes d'intelligence artificielle.

Sources :

IBM. *Comprendre les différents types d'intelligence artificielle.*

CNIL. *Apprentissage automatique (Machine Learning).*

Fortinet. *L'intelligence artificielle (IA) dans la cybersécurité.*